FETAL-MATERNAL MONITORING IN THE AGE OF ARTIFICIAL INTELLIGENCE AND COMPUTER-AIDED DECISION SUPPORT: A MULTIDISCIPLINARY PERSPECTIVE

EDITED BY: Antoniya Georgieva, Patrice Abry, Martin Gerbert Frasch and Ines Maria Nunes

PUBLISHED IN: Frontiers in Pediatrics, Frontiers in Artificial Intelligence and Frontiers in Medicine

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-239-6
DOI 10.3389/978-2-83250-239-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

FETAL-MATERNAL MONITORING IN THE AGE OF ARTIFICIAL INTELLIGENCE AND COMPUTER-AIDED DECISION SUPPORT: A MULTIDISCIPLINARY PERSPECTIVE

Topic Editors:
Antoniya Georgieva, University of Oxford, United Kingdom
Patrice Abry, Université de Lyon, France
Martin Gerbert Frasch, University of Washington, United States
Ines Maria Nunes, North Maternal Child Center Albino Aroso (CMIN), Portugal

Citation: Georgieva, A., Abry, P., Frasch, M. G., Nunes, I. M., eds. (2022). Fetal-Maternal Monitoring in the Age of Artificial Intelligence and Computer-Aided Decision Support: A Multidisciplinary Perspective. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-239-6

Table of Contents

frontiers | Frontiers in Pediatrics

TYPE Editorial
PUBLISHED 05 September 2022
DOI 10.3389/fped.2022.1007799

OPEN ACCESS

EDITED AND REVIEWED BY
Hannes Sallmon,
Deutsches Herzzentrum
Berlin, Germany

*CORRESPONDENCE
Martin G. Frasch
mfrasch@uw.edu

SPECIALTY SECTION
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

RECEIVED 30 July 2022
ACCEPTED 22 August 2022
PUBLISHED 05 September 2022

CITATION
Georgieva A, Abry P, Nunes I and Frasch MG (2022) Editorial: Fetal-maternal monitoring in the age of artificial intelligence and computer-aided decision support: A multidisciplinary perspective. *Front. Pediatr.* 10:1007799. doi: 10.3389/fped.2022.1007799

Editorial: Fetal-maternal monitoring in the age of artificial intelligence and computer-aided decision support: A multidisciplinary perspective

Antoniya Georgieva[1], Patrice Abry[2], Ines Nunes[3,4] and Martin G. Frasch[5]*

[1]Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, United Kingdom, [2]CNRS, École Normale Supérieure de Lyon, Laboratoire de Physique, Lyon, France, [3]Centro Materno Infantil Do Norte–Centro Hospitalar Universitário Do Porto, Porto, Portugal, [4]Centro Académico Clínico, Instituto de Ciências Biomédicas Abel Salazar, Centre for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal, [5]Department of Obstetrics and Gynecology and Center on Human Development and Disability, University of Washington, Seattle, WA, United States

KEYWORDS

electronic fetal monitoring, cardiotocography, electrocardiography, electroencephalography, pregnancy, machine learning, decision support, big data and analytics

Editorial on the Research Topic
Fetal-maternal monitoring in the age of artificial intelligence and computer-aided decision support: A multidisciplinary perspective

Introduction

Across the globe, each day, we continue to have term babies arrive at delivery wards in good condition *in utero*, only to be born hours later with neurological injuries (1). The consequences are profound and life-long for the babies, parents, siblings, and their wider family (2). Clinical staff involved in the obstetric management are severely impacted in multiple ways. On the other hand, Cesarean section to avoid oxygen deprivation during labor carries multiple risks for mother, fetus, future pregnancies; as well as costs. But achieving safe spontaneous delivery is sometimes challenging due to poorly understood and complex fetal physiology, and often, conflicting healthcare needs for mother and baby.

In developed countries, the standard of care for pregnancies deemed at risk is continuous electronic fetal monitoring with cardiotocography (CTG) during

labor—displaying fetal heart rate (FHR) along uterine contractions on a long paper strip, typically assessed by eye. CTG interpretation continues to be a massive challenge with high false positive rate and poor sensitivity (3). The evidence how to interpret CTG traces is relatively limited, derived by subjective clinical experience and by animal studies, which have own limitations, but provide invaluable core evidence (4–7). For example, there has been an ongoing concern about the persistent use of "increased" or "decreased" heart rate variability (HRV) which goes against the evidence of what variability represents—a complex dynamic multi-dimensional pattern, such that a single-order description of this pattern in terms of "ups" and "downs" discards most of its predictive information (8). Furthermore, the CTG technology itself is imperfect with frequent signal loss, noise, and confusion between maternal and fetal heart rates (3).

So, the main challenge remains lack of technology to monitor reliably the fetus *in utero*, at a time when labor brings unprecedented challenges to fetal oxygen supply and forces the fetus to rely on compensatory mechanisms and reserves. Consequently, the adverse outcomes we wish to prevent are both heterogeneous and rare. This means that, to develop new detection/prediction methods or algorithms based on the CTG and utilize our modern computing and data science capabilities, we need to obtain CTG and maternity data at scale (9). And there is a need of "feature engineering", i.e., bespoke signal processing methods to account for the noisy, low sampling rate nature of the signal (10). Despite these known shortcomings and promising opportunities to make considerable impact on human health, improving fetal monitoring through novel technologies continues to be a niche field, especially for large scale clinical use. It remains unclear how to best apply computers and large datasets for clinical benefit, hand-in-hand with novel engineering solutions.

Therefore, we assembled this first of its kind Frontiers Research Topic, focused on multidisciplinary intrapartum risk assessment through technology and clinical insights. It builds on our experience and existing collaboration in organizing the bi-annual international workshop—Signal Processing and Monitoring in Labor—providing multidisciplinary forum for the clinical and engineering challenges of fetal monitoring during labor. These workshops have included experts from academia and industry representing multidisciplinary domains of clinical medicine (obstetrics, neonatology), physiology, physics, epidemiology, data sciences, statistical signal processing, artificial intelligence (AI), and signal feature and software engineering. We hope that the multi- and transdisciplinary character of these workshops can serve as a template for the framework in which solutions to the problem of fetal monitoring intrapartum can be found. In the following, we synthesize the core insights provided by the 15 contributions (four systematic reviews, one opinion and ten original research articles) of this Research Topic. We then discuss the future directions for this field.

Systematic reviews

O'Sullivan et al. review the decision support systems used in three RCTs for intrapartum CTG, summarizing the algorithms, the outcomes of the trials and the limitations (O'Sullivan et al.). Preliminary work suggests that the inclusion of clinical data can improve the performance of AI-assisted CTG. Combined with newer approaches to the classification of CTG traces, this offers promise for rewarding future development.

Castel et al. screened 256 studies in four languages and arrived at 40 studies in the qualitative and quantitative analysis of the intrapartum fetal electroencephalogram (fEEG). The authors show its potential to act as a direct biomarker of fetal brain health during delivery, ancillary to FHR monitoring and readily feasible using the presently used fetal scalp electrode. Real world evidence of fetal EEG acquired from a regular fetal scalp electrode is also presented. Highlighted is the need for clinical prospective studies to further establish the utility of intrapartum fEEG monitoring intrapartum, suggesting suitable clinical study designs.

Ribeiro et al. focused on non-linear analysis of FHR based on concepts of chaos, fractality, and complexity: entropies, compression, fractal analysis, and wavelets. The authors aim to increase our knowledge about cardiovascular dynamics in healthy and pathological fetuses. Two hundred and seventy articles are included in the review. The top five primary research objectives covered by the selected papers are detection of hypoxia, maturation or gestational age, intrauterine growth restriction, and fetal distress.

This review shows that non-linear indices can be used but are not yet applied in clinical practice. Some studies show that the combination of several linear and non-linear indices would be ideal for improving the analysis of the fetal wellbeing. Future studies should narrow the research question so a meta-analysis could be performed, probing the indices' performance.

Castro et al. review the spectral bands reported in intrapartum FHR studies and evaluate their performance in detecting fetal acidemia. Twenty-five (out of 176) studies are included. An open-access FHR database is used, with recordings of the last half an hour of labor of 246 fetuses. Four different umbilical artery pH cut-offs are considered for fetuses' classification into acidemic or non-acidemic: 7.05, 7.10, 7.15, and 7.20. The area under the receiver operating characteristic curve (AUROC) is used to quantify the frequency bands' ability to distinguish acidemic fetuses.

Bands referring to low frequencies, mainly thought to be associated with neural sympathetic activity, are found to be the best at detecting acidemic fetuses, with the more severe

definition (pH $\leq$ 7.05) attaining the highest values for the area under the ROC–AUROC [0.770 (0.608–0.932)].

This study shows that the power spectrum analysis of the FHR is a simple and powerful tool that has potential for CTG evaluation and helping healthcare professionals to accurately identify fetuses at risk of acidemia. Naturally, prospective clinical evaluations are needed. Despite the clear potential of many HRV metrics to predict fetal acidemia, the proxy of fetal acidemia itself has been proven to not be appropriate for predicting clinical outcome (11, 12). As such, while to date, most studies in the field have sought to predict pH at birth, some studies in this Research Topic and elsewhere have sought to predict physiologically more direct outcomes related to fetal compromise, either using FHR data as time series or as scanned CTG tracings (Gold et al.; Roux et al.).

In this context, the opinion paper by Schifrin provides insights into the complexities of CTG interpretation that transcend the conventional scientific reasoning applied in this field and extend to historical and medicolegal confines as well as epistemological biases. The author emphasizes the problem of the poor definition of outcome, centered on pH prediction and avoidance of low pH while providing little guidance to intrapartum management. He discusses opportunities for future research which, in part, are tackled in the above-mentioned reviews, as well as in the research articles of this topic we discuss in the following paragraphs.

Original contributions

It has been striking to observe that all studies in the present Research Topic relied on retrospective data analysis. It highlights the major issue in our field which impedes rapid innovation: the data required to test models predicting adverse outcome with clinically actionable performance must be large (millions of CTGs) and no single institution, company or research team have access to such data at this time.

Several studies deal with the signal quality issues which also hamper the progress in the field. While some approaches seek to improve the ultrasound-based CTG signal and its derived HRV estimates by thoughtfully engineering signal processing and machine learning techniques (Roux et al.; Vargas-Calixto et al.), others focus on the emerging technologies such as transabdominal ECG (Fotiadou and Vullings; Vullings and van Laar). It is worth noting that both EEG, reviewed by Castel et al., and transabdominal ECG technologies have been studied for at least 80 and 40 years, respectively. However, these technologies have not yet found a broad clinical adoption.

Pursuing further the issue of FHR signal quality in conventional CTG, Vullings and van Laar conduct a quantitative comparison of FHR derived from two commercial intrapartum ECG-based fetal monitors, one using fetal scalp electrode and another using transabdominal ECG sensors. The authors compare the FHR detection rates to those from a conventional ultrasound-based CTG. They report a reliable FHR in >95% of time intrapartum which represents a substantial improvement over Doppler ultrasound. During second stage of labor, given stronger contraction and FHR drops to below 100 bpm, the signal processing challenges remain considerable and the performance of the method decreases. However, with a reliability higher than 80%, the proposed method still outperforms Doppler ultrasound and other reference methods by a significant amount.

Vargas-Calixto et al. deploy a signal processing approach to identify robustness of select HRV estimates to noise contained in conventional ultrasound-based FHR. Such insights hold promise of improved fidelity of HRV estimation given the constraints of CTG signal in terms of quality of beat-to-beat estimation and noisiness. Their work highlights the importance of considering the nature of the underlying FHR signal when selecting and trusting the HRV metrics derived from the signal.

Fotiadou and Vullings turn their attention to the alternative approach of FHR derivation promising superior signal quality and patient experience compared to conventional ultrasound methods, the transabdominal ECG recorded antepartum and intrapartum. The authors present a method to extract the entire multi-channel fetal ECG waveform using deep convolutional neural networks (CNN), a broadly used deep learning technique for tasks such as image analysis and here showing promise for maternal-fetal ECG deconvolution. Meanwhile, in Frasch et al., CNN was also used to detect pathognomonic CTG patterns directly from images. This highlights the broad potential of deep learning techniques in various settings of fetal monitoring.

Costa et al. report the performance of their intrapartum CTG system SisPorto, focusing on prediction of pathological FHR patterns from hitherto underutilized HRV fragmentation metrics using the open-source Brno/Prague CTG database. In another retrospective study, Lovers et al. used computerized methods to analyze CTGs from ~28,000 births and identify presence of abnormalities in the first hour CTG as well as associated clinical risk factors. This highlights the importance of admission CTG analysis for labor management/triage. Another facet of this study is the indication of the importance of antepartum fetal health that likely precipitates the abnormalities seen in the admission CTG. The creation of maternal-fetal monitoring technologies that track health of mother-fetus dyad antepartum is much needed and subject of ongoing research and development.

In addition, Pini et al. contribute to the detection of late intrauterine growth restriction (IUGR) in a retrospective case-control design at 38 weeks of gestation, on a publicly available dataset of CTGs (Pini et al.). The HRV feature engineering the authors present for their machine learning model accounts for the CTG's properties such as signal duration and signal quality. Future studies should attempt to validate these findings in larger datasets and in admission

CTGs to mimic the clinical scenario of intrapartum triaging at admission.

In their complementary work, Gold et al. and Roux et al. study a fetal sheep model-based FHR dataset of umbilical cord occlusions (UCO) of increasing severity mimicking the uterine contractions as they may occur during first and second stage of human labor. Both teams engineer novel machine learning techniques for prediction of a physiologically and clinically meaningful outcome of fetal cardiovascular decompensation, rather than acidemia, and to address the complex nature of fetal HRV signal accounting for the constraints on signal quality in terms of sampling rate and noise. Both studies present novel individualized approach to machine learning of FHR data. This is a promising avenue to exploit in future studies, even when the data size is relatively small, as it leverages individual variability at different time scales, to identify departure from the "normal" phenotype. And, using the same experimentally derived dataset, Rivolta et al. present novel HRV features of deceleration reserve as a distinguishing property of chronically hypoxic fetuses.

Finally, in their study on scanned CTG recordings, Frasch et al., for the first time, deploy computer vision techniques of deep learning, a form of AI, to identify important patterns of CTG on images, rather than the raw data. Importantly, the authors open-sourced their algorithm including the approach to annotate online the CTG data. This could potentially lead to better collaboration, for example crowdsourced CTG annotation.

Conclusions and outlook for the future

Intrapartum fetal monitoring is, hopefully, on the verge of technological disruption thanks to the recent advances in and convergence of computing resources on edge and in the cloud, AI and the resulting emergence of digital health as a field. The clinical need remains unmet and, more than ever the chasm, between technological advances and possibilities and the reality of fetal monitoring around the world is wide and asking for closure. Thankfully, the technological disruption through innovative devices and algorithms is being pursued in Europe and USA as we have seen from the academic and industry partners in this Research Topic, as well as in the relevant Signal Processing and Monitoring Workshops in 2019 (Porto, Portugal) and the most recent 2022 workshop in Munich (Germany). There is a strong consensus across the board that

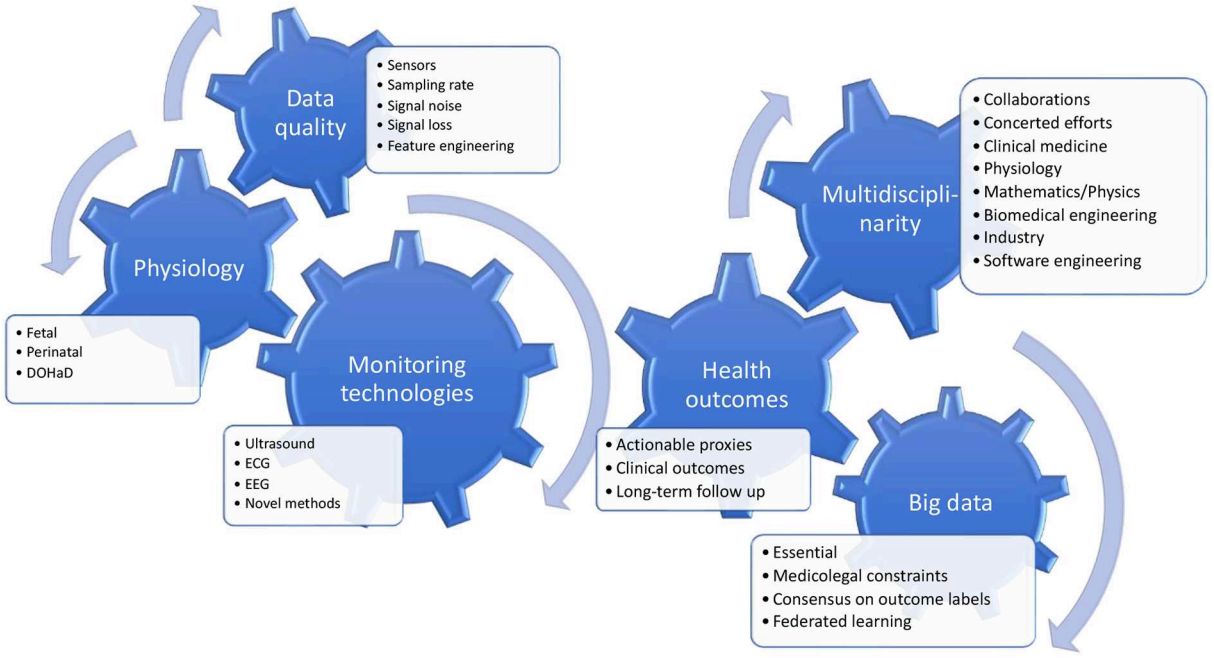

FIGURE 1
Key insights from the Research Topic and future directions. There is a growing awareness of the antecedents of intrapartum fetal reserve for the trial of labor which require an integration of the physiology of whole pregnancy and the well-known relationships between intrauterine adversity on one hand, and the perinatal and postnatal developmental trajectories on the other hand (Developmental Origins of Health and Disease, the DOHaD concept). Another key insight is the requirement for clinically actionable outcome labels in the prediction models that are being developed and the recognition of the fundamental constraints on the data sizes of the individually accessible cohorts. Therefore, there is a clear need for multinational and multidisciplinary work to address the different challenges and research questions, which are all integral to successfully improving the technologies for intrapartum fetal monitoring.

the challenge of intrapartum fetal monitoring can only be solved through collective, multinational effort. The big data (millions of CTGs) required is impossible to collect by any single stakeholder. In parallel, physiological research needs to continue to address the fundamental questions raised in this Research Topic and elsewhere about the mechanisms of injury. Also required is the push for innovative technologies that can acquire other important signals for the fetus, i.e., ECG, EEG, and beyond.

As the challenge is enormous, for a proposed shared partnership effort to succeed, we need to form a cohesive view of the shared direction. In Figure 1, we present such a cohesive view of the shared and interacting priorities, representing also our main take-home message from this Research Topic.

Author contributions

AG and MF prepared the first draft of the manuscript. All authors edited and approved the final version.

Funding

Author AG was funded by the UK National Institute of Health and Care Research (CDF-2016-09-004 and NIHR202117).

Acknowledgments

The authors acknowledge the enthusiastic and dedicated contributions of all participants of the Signal Processing and Monitoring Workshops in 2019 (Porto, Portugal) and the most recent 2022 workshop in Munich (Germany). The numerous inputs of all participating research teams from academia and industry are what enabled this Research Topic and resulted in the presented evidence and inference for future directions.

Conflict of interest

Author MF holds patents on EEG and ECG processing, the founder of and consults for Digital Health companies commercializing the predictive potential of physiological time series for human health.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

References

1. Gale DC, Stanikov ME, Jawad S, Uthaya DS, Modi PN. *Brain Injury Occurring During or Soon After Birth: A Report for the National Maternity Ambition Commissioned by the Department of Health*. Imperial College London, UK (2018).

2. Lee BL, Glass HC. Cognitive outcomes in late childhood and adolescence of neonatal hypoxic-ischemic encephalopathy. *Clin Exp Pediatr.* (2021) 64:608–18. doi: 10.3345/cep.2021.00164

3. Alfirevic Z, Gyte GM, Cuthbert A, Devane D. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Datab Syst Rev.* (2017) CD006066. doi: 10.1002/14651858.CD006066.pub3

4. Ayres-de-Campos D, Spong CY, Chandraharan E, and FIGO Intrapartum Fetal Monitoring Expert Consensus Panel. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynaecol Obstet.* (2015) 131:13–24. doi: 10.1016/j.ijgo.2015.06.020

5. Lear CA, Galinsky R, Wassink G, Yamaguchi K, Davidson JO, Westgate JA, et al. The myths and physiology surrounding intrapartum decelerations: the critical role of the peripheral chemoreflex. *J. Physiol.* (2016) 594:4711–25. doi: 10.1113/JP271205

6. NICE. *Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth*. NICE (2017).

7. Westgate JA, Wibbens B, Bennet L, Wassink G, Parer JT, Gunn AJ. The intrapartum deceleration in center stage: a physiologic approach to the interpretation of fetal heart rate changes in labor. *Am J Obstetr Gynecol.* (2007) 236.e1–11. doi: 10.1016/j.ajog.2007.03.063

8. Frasch MG. Saving the brain one heartbeat at a time. *J Physiol.* (2018) 596:5503–4. doi: 10.1113/JP275776

9. Georgieva A, Abry P, Chudáček V, Djurić PM, Frasch MG, Kok R, et al. Computer-based intrapartum fetal monitoring and beyond: a review of the 2nd workshop on Signal Processing and Monitoring in Labor (October 2017, Oxford, UK). *Acta Obstet Gynecol Scand.* (2019) 98:1207–17. doi: 10.1111/aogs.13639

10. Frasch MG. Sampling rate and heart rate variability: on metrics and health outcomes. *J Biomed Inform.* (2022) 129:104061. doi: 10.1016/j.jbi.2022.104061

11. 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) 168:155–60. doi: 10.1016/j.ejogrb.2013.01.003

12. Gonen N, Gluck O, Zussman NM, Bar J, Kovo M, Weiner E. The role of umbilical cord gas studies in the prediction of adverse neonatal outcomes in scheduled nonlaboring term singleton cesarean deliveries. *Am J Obstetr Gynecol.* (2019) 1:119–27. doi: 10.1016/j.ajogmf.2019.06.001

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 26 August 2020
doi: 10.3389/fped.2020.00508

Multi-Channel Fetal ECG Denoising With Deep Convolutional Neural Networks

Eleni Fotiadou and Rik Vullings*

Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands

Non-invasive fetal electrocardiography represents a valuable alternative continuous fetal monitoring method that has recently received considerable attention in assessing fetal health. However, the non-invasive fetal electrocardiogram (ECG) is typically severely contaminated by a considerable amount of various noise sources, rendering fetal ECG denoising a very challenging task. This work employs a deep learning approach for removing the residual noise from multi-channel fetal ECG after the maternal ECG has been suppressed. We propose a deep convolutional encoder-decoder network with symmetric skip-layer connections, learning end-to-end mappings from noise-corrupted fetal ECG signals to clean ones. Experiments on simulated data show an average signal-to-noise ratio (SNR) improvement of 9.5 dB for fetal ECG signals with input SNR ranging between −20 and 20 dB. The method is additionally evaluated on a large set of real signals, demonstrating that it can provide significant quality improvement of the noisy fetal ECG signals. We further show that employment of multi-channel signal information by the network provides superior and more reliable performance as opposed to its single-channel network counterpart. The presented method is able to preserve beat-to-beat morphological variations and does not require any prior information on the power spectra of the noise or the pulse location.

Keywords: convolutional neural networks, encoder-decoder network, fetal ECG denoising, fetal ECG enhancement, fetal electrocardiography

OPEN ACCESS

Edited by:
Antoniya Georgieva,
University of Oxford, United Kingdom

Reviewed by:
Fernando Andreotti,
University of Oxford, United Kingdom
Petar M. Djuric,
Stony Brook University, United States

****Correspondence:***
Eleni Fotiadou
e.fotiadou@tue.nl

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *08 May 2020*
Accepted: *17 July 2020*
Published: *26 August 2020*

Citation:
Fotiadou E and Vullings R (2020) Multi-Channel Fetal ECG Denoising With Deep Convolutional Neural Networks. Front. Pediatr. 8:508. doi: 10.3389/fped.2020.00508

INTRODUCTION

The fetal electrocardiogram (ECG) can be used to monitor the condition of the fetal heart from early pregnancy until delivery (1). Nowadays, fetal monitoring is mainly performed by cardiotocography or by ECG recordings where an electrode is directly placed on the fetal scalp. Cardiotocography records the fetal heart rate together with the uterine contractions. The advantages of the method are that it is performed non-invasively and is safe for the patient. On the other hand, it is prone to signal loss, while recorded changes of the heart rate are not always precise (2). Scalp ECG recordings are a more reliable means of monitoring the fetal health. However, they are invasive, may pose a health risk to the fetus, and can only be performed during labor, when the membranes have ruptured.

Non-invasive fetal electrocardiography, performed by placing electrodes on the maternal abdomen, is a promising alternative to standard fetal monitoring. In comparison with cardiotocography, it provides more accurate information because it does not need to average over multiple beats for the heart rate extraction. Moreover, it provides the possibility to assess the ECG

morphology, related to the electrical activity of the fetal heart. The advantage of the method over the scalp ECG measurements is that it can be performed already during pregnancy, it is safe for the fetus and comfortable for the mother. However, the difficulty to extract a clean fetal ECG from the abdominal mixture is the main reason that the application of the method in clinical practice is still limited. The interferences and noises in the abdominal recordings among others include the maternal ECG, powerline interference, baseline wander, muscle noise from the fetus and mother and movement artifacts. Considering that the signals of some of these interferences overlap both in time and frequency with the fetal ECG, the extracted fetal ECG signals usually have very low signal-to-noise ratio (SNR). Therefore, the non-invasive recordings are in practice merely used for heart rate analysis.

There are typically three main steps in the fetal ECG extraction process; preprocessing, separation and postprocessing (3). Preprocessing includes removal of unwanted noise such as powerline interference and baseline wander. In the separation step, the maternal ECG is estimated and then subtracted from the signals to obtain the fetal ECG. Finally postprocessing is employed to enhance the quality of the extracted fetal ECG signals. The work on non-invasive fetal ECG analysis has mainly targeted the first two steps, together with the improvement of the acquisition devices (4), while only few works focused on the postprocessing of the obtained signals. Beat-to-beat averaging is a traditional method which is often used to improve the SNR of the extracted signals, at the expense of losing individual variations in pulse shape (5). Different wavelet denoising techniques were additionally proposed in the literature for the postprocessing of the extracted fetal ECG signals (6, 7). In a previous work (8), the authors employed an augmented time-sequenced adaptive filter to enhance the quality of the extracted fetal ECG. Despite the significant quality improvement that the method achieves, the location of the fetal pulses is required to synchronize the filter and the method cannot handle abrupt changes in fetal ECG morphology, e.g., in cases of arrhythmia.

Recently, deep neural network models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs) and stacked denoising autoencoders have been successfully applied for a variety of purposes including signal and image denoising (9–13). Moreover, few works reported adult ECG signal denoising (14, 15), fetal QRS detection (16, 17), and fetal ECG signal reconstruction (18). Zhong et al. (19) presented a deep convolutional encoder-decoder framework for preprocessing abdominal recordings to remove noise. However, they did not extract the fetal ECG from the preprocessed signals to ensure that it is not suppressed by the network. The authors were the first to propose a deep convolutional encoder-decoder network for postprocessing non-invasive single-channel fetal ECG (20, 21), achieving a substantial quality improvement of the noisy signals. The method tackled some of the shortcomings of the state-of-the-art non-invasive fetal ECG postprocessing methods, since it can preserve beat-to-beat morphological variations and does not require prior knowledge about the location of the fetal pulses. However, in cases of heavily corrupted signals, the method was unable to reliably reconstruct some relevant morphological features of the ECG, sporadically even causing presence of "fake" waves, i.e., waves in the reconstructed ECG that should not have been there or should have had opposite sign. For a practical application this might be dangerous, leading to wrong diagnosis.

In this work, we are dealing with the aforementioned problem by extending our model to handle multiple fetal ECG channels. Multiple electrodes measure the electrical activity of the heart from different angles. We propose to use a deep convolutional encoder-decoder network with symmetric skip connections that learns how to optimally combine the input channels to deliver a reliable clean, multi-channel ECG as output. The method eliminates the residual noise in the fetal ECG by capturing the signal structure in the convolutional layers and recovering the details by the transposed convolutional layers.

MATERIALS AND METHODS

Data

Simulated Data

For the training, but also for the evaluation of the proposed network, we created an extensive simulated fetal ECG dataset that consists of two parts. The first part was built by employing the fecgsyn toolbox developed by Behar et al. (22, 23). The toolbox enables the creation of abdominal mixtures with adjustable noise sources, heart rate, heart rate variability, fetal movement, ectopic beats and contractions. A Gaussian model is used to simulate the ECG beats, as originally developed by McSharry (24) and further improved by Sameni (25). Any number of electrodes can be positioned on the maternal abdomen for the simulations. Unfortunately, the simulated fetal ECGs are based merely on 9 available vectorcardiograms (VCGs). Since there is limited variation in the shape and lengths of the individual PQRST waves in these VCGs, there is an increased risk of overfitting the network. This means that the network might learn to reproduce these limited morphologies and enforce resemblance of the denoised signals with the training data. In fact, what happened in our initial experiments is that the P and T waves of the denoised signals were shifted with respect to their ground truth data to match the locations of the training data. For this reason, we built a modified version of the toolbox that creates a variety of new ECG morphologies based on the already available VCGs. The modified toolbox receives a VCG as input, alters the length of the VCG intervals along with the amplitudes of the PQRST waves and subsequently uses it as a base to form the abdominal fetal ECG. Initially, for all 9 VCGs, the points of interest, which are the beginning and end of the P wave, T wave, and QRS complex were annotated and saved to be later available to the simulator. In every iteration of the modified simulator, one of the 9 VCGs is randomly selected and subsequently the start and the end of the waves are randomly shifted in position. Since the shift of the start and shift of the end point of each wave are not identical, also the length of the waves is automatically varied this way. The amplitude of each wave is changed as well by random scaling. The modified VCG is the starting point that the abdominal fetal ECG can be created. With the help of the modified toolbox we created a large dataset of four-channel abdominal mixtures, where different physiological events were considered, such as heart rate decelerations and accelerations,

fetal movement, ectopic beats, uterine contraction etc., similar to the Fetal ECG Synthetic Database (26). The VCG alterations were chosen so as to include an ample range of variations of the ECG morphological features, while still ensuring their physiological plausibility. When obtaining the simulated data, we varied the placement of the four electrodes to make the method invariant to variations in the electrode position.

To further enrich the ECG morphologies in our dataset and reduce the risk of overfitting to the training data we generated an additional set of simulated signals based on adult ECG from the PTB Diagnostic ECG Database of Physionet (27). The database comprises of both normal and pathological signals with 15 leads, sampled at 1,000 Hz. 549 records from 290 male and female subjects are available. Adult and fetal ECG have similar morphology but the adult ECG intervals and amplitudes are larger compared to the fetus. The adult ECG was preprocessed to remove noise and resemble the fetal ECG. First, a high-pass filter with cut-off frequency of 1 Hz was applied followed by Savitzky-Golay filtering of order 8 and length 31. Afterwards, considering that the fetal heart beats two to three times faster than the adult heart, the signals were resampled to half frequency. Adjustment of the signals amplitude was not necessary because they were, in a later data preparation step, anyway normalized before entering the network. As a next step, four-channel signals were created by making random combinations of four leads, where a maximum of two was chosen out of the six first limb leads. Finally, "real" noise was added to the signals. For the "real" noise we employed a set of six-channel abdominal recordings of an ongoing study of which the protocol is described in (28). In a subset of these recordings we found it impossible to detect the fetal ECG, either because of the shielding of the fetus by the vernix caseosa or because the fetal heart was far from some electrodes. We considered that these measurements, after the maternal ECG suppression and powerline interference removal, consist of pure noise and added them to the preprocessed adult ECG to generate our simulated fetal ECG signals.

Real Data

In order to evaluate how well our algorithm performs in real signals we employed two databases. The first one is a private set of non-invasive fetal ECG measurements, obtained in collaboration with the Máxima Medical Center, Veldhoven, the Netherlands (28, 29). The dataset contains 462 six-channel recordings of different women, at least 18 years old, between 18 and 24 weeks of gestation. The fetal ECG was recorded with adhesive Ag/AgCI electrodes on the abdomen of the pregnant women while they were in semi-upright position. Six electrodes were placed around the navel to produce six channels of electrophysiological measurements, while two additional electrodes, placed close to the navel, served as common reference and ground. Each recording lasted from 5 up to 50 min and was digitized and stored at 500 Hz sampling frequency by a fetal monitoring system (Nemo Healthcare BV, The Netherlands). Since the signals were measured through six electrodes, we selected the first, third, fourth and fifth dimensions to form the four-channel fetal ECG signal.

The second real dataset is the Abdominal and Direct Fetal Electrocardiogram Database which consists of four-channel abdominal fetal ECG recordings obtained by five women in labor, between 38 and 41 weeks of gestation (30). Each recording comprises four different signals acquired from the maternal abdomen together with a reference direct fetal ECG registered from the fetal head. The configuration of the abdominal electrodes consisted of four electrodes placed around the navel, a reference electrode placed above the pubic symphysis and a common reference electrode placed on the left leg. The recordings have duration of 5 min and are sampled at 1,000 Hz.

Data Preprocessing

The signals of all the datasets were preprocessed before entering the network. The fetal ECG extraction was performed with the help of the open-source algorithm of Varanini et al. (31) and the signals were resampled to 500 Hz to have a common reference. Finally, the fetal ECG signals were divided in segments of 1920 $\times$ 4 samples and normalized to have zero mean and unity standard deviation. The normalization was performed along each channel separately.

Network Description

The proposed fetal ECG denoising CNN network is illustrated in **Figure 1**. It consists of an encoder of eight convolutional layers and a decoder of eight symmetric transposed convolutional layers. The network receives a noisy fetal ECG signal as input and delivers a denoised one as output. The convolutional layers act as a feature extractor which captures the abstraction of the fetal ECG while eliminating the noise. Subsequently, the transposed convolutional layers decode the fetal ECG abstraction to recover the signal details. The convolutional layers are symmetrically connected with the transposed convolutional ones via skip connections. The role of the skip connections is two-fold. First, they help back-propagating the gradients to bottom layers, facilitating the training of our deep network. Second, they pass signal content from the bottom to top layers to aid in recovering the signal details.

The non-invasive fetal ECG typically contains a high amount of noise and thus a large denoising patch can lead to more efficient noise removal by using context information from a larger signal region. It was indicated in the literature that the denoising patch is highly correlated with the receptive field of the network, i.e., the region in the input space that a CNN feature can be affected by (11, 32). The receptive field of the network is determined by the kernel size, the depth of the network and whether subsampling or dilation is used in the convolution operations. A common approach to increase the receptive field is to increase the number of layers in the network but this is computationally expensive. We chose to use a relatively deep network of eight convolutional and eight transposed convolutional layers. Since our data are temporal, we adopt one-dimensional convolutions and transposed convolutions. In addition, subsampling by two is performed after each convolutional layer, apart from the first, and upsampling by two after the transposed convolutional layers, apart from the last

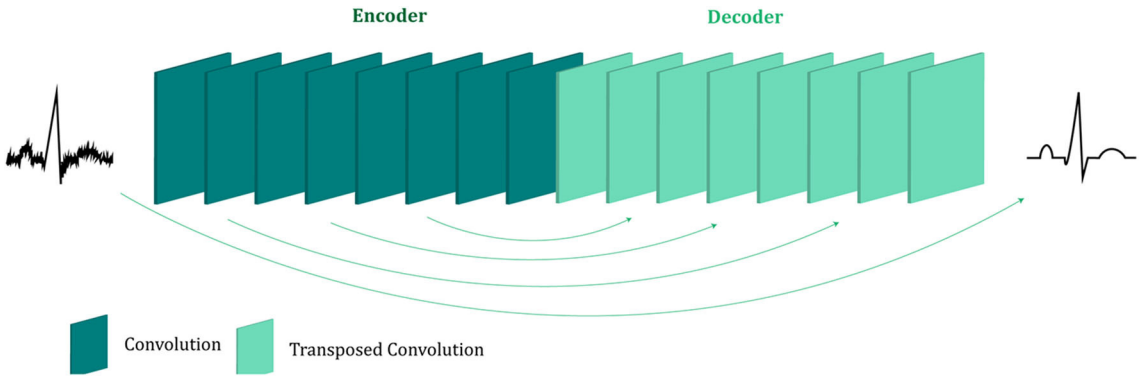

FIGURE 1 | The architecture of the proposed multi-channel fetal ECG denoising network. The network consists of an encoder of eight convolutional layers and a decoder of eight transposed convolutional layers, which are linked symmetrically by skip connections.

one. Subsampling operations are not generally preferred in denoising tasks in order to preserve the signal details (10). On the other hand, in our case they lead to a significant increase of the receptive field, necessary for removing the large amount of noise present in the fetal ECG signals. Moreover, in order to exploit the self-similarity of the ECG signals the network should permit the convolutions to extend to several heartbeats. Regarding the kernel size we empirically determined that 15 achieves satisfactory results by being large enough to include sufficient signal information without excessively increasing the number of network parameters. The input and output of the network have dimension 1920 $\times$ 4 which corresponds to four-channel ECG of 3.84 s. For non-linearity after each layer, leaky rectified linear units (LeakyRelu) with a slope of 0.2 are utilized. The aforementioned parameter choices led to a receptive field of roughly 4 s that corresponds to 5–10 heartbeats. A detailed description of the network architecture and the parameters is given in **Table 1**.

Skip Connections

In shallow networks transposed convolutions works well for recovering the signal details but as the network goes deeper, they do not longer work satisfactory (9). Our network is deep and heavy subsampling is performed for the sake of increasing the receptive field of the network, resulting in significant loss of signal information. To address this issue, skip connections are added between every two convolutional and mirrored transposed convolutional layers as shown by the arrows in **Figure 1**. The skip connections carry signal information and account to a great extent for the lost signal details introduced by the subsampling. Moreover, these skip connections allow the gradient update rules to back-propagate to the bottom layers directly, dealing with the gradient vanishing problem occurring in deep architectures. The way that the skip connections are used in the network is depicted in **Figure 2**.

Network Training

For training the network the normalized mean squared error loss was minimized, which is defined as:

$$\mathcal{L} = \frac{1}{N*L*M} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{m=1}^{M} \frac{\left(X_{clean\,n,l,m} - X_{denoised\,n,l,m}\right)^2}{\overline{X}^2_{clean\,n,l}}, \tag{1}$$

where N is the number of the training data in a batch, L is the number of channels, M is the length of the signals, X represents the fetal ECG and $\overline{X}^2$ is the mean squared amplitude of X. In our experiments N = 64, L = 4 and M = 1920. The Adam algorithm was selected (33) as an optimization algorithm while the learning rate was set to 0.00001. The training method that we followed is supervised, meaning that we need clean fetal ECG signals as labels together with the noisy signals. For this reason, the training of the network was performed based only on simulated data. The simulated data were separated in two sets for the training and testing of the method. The training set contains the signals simulated by the modified fecgsyn toolbox based on VCG 1-7 and 449 preprocessed records from 212 subjects of the PTB dataset. The test set contains the simulated signals based on VCG 8-9 from the modified fecgsyn toolbox, plus 100 preprocessed records of 78 subjects of the PTB dataset. The SNR of the training set ranges from −15 to 15 dB. The network was trained for 21 epochs until convergence was reached.

Performance Evaluation

In the simulated dataset, the performance of the method was evaluated based on the SNR improvement of the fetal ECG signals achieved by the network. The metric is estimated for a channel, l, of a signal as:

$$SNR_{imp} = 10log_{10} \frac{\sum_{m=1}^{M} \left|X_{noisy_{l,m}} - X_{clean\,l,m}\right|^2}{\sum_{m=1}^{M} \left|X_{denoised\,l,m} - X_{clean\,l,m}\right|^2}. \tag{2}$$

TABLE 1 | Detailed overview of the proposed network architecture.

	Layer	Output size	Filter size	Kernel size
Encoder	Convolution (stride = 1)	1920 × 64	64	15
	LeakyRelu(0.2)	1920 × 64		
	Convolution (stride = 2)	960 × 128	128	15
	LeakyRelu(0.2)	960 × 128		
	Convolution (stride = 2)	480 × 256	256	15
	LeakyRelu(0.2)	480 × 256		
	Convolution (stride = 2)	240 × 256	256	15
	LeakyRelu(0.2)	240 × 256		
	Convolution (stride = 2)	120 × 512	512	15
	LeakyRelu(0.2)	120 × 512		
	Convolution (stride = 2)	60 × 512	512	15
	LeakyRelu(0.2)	60 × 512		
	Convolution (stride = 2)	30 × 1024	1024	15
	LeakyRelu(0.2)	30 × 1024		
	Convolution (stride = 2)	15 × 2048	2048	15
	LeakyRelu(0.2)	15 × 2048		
Decoder	Transposed Convolution(stride = 2)	30 × 1024	1024	15
	LeakyRelu(0.2)	30 × 1024		
	Transposed Convolution(stride = 2)	60 × 512	512	15
	Addition	60 × 512		
	LeakyRelu(0.2)	60 × 512		
	Transposed Convolution(stride = 2)	120 × 512	512	15
	LeakyRelu(0.2)	120 × 512		
	Transposed Convolution(stride = 2)	240 × 256	256	15
	Addition	240 × 256		
	LeakyRelu(0.2)	240 × 256		
	Transposed Convolution(stride = 2)	480 × 256	256	15
	LeakyRelu(0.2)	480 × 256		
	Transposed Convolution(stride = 2)	960 × 128	128	15
	Addition	960 × 128		
	LeakyRelu(0.2)	960 × 128		
	Transposed Convolution(stride = 2)	1920 × 64	64	15
	LeakyRelu(0.2)	1920 × 64		
	Transposed Convolution(stride = 1)	1920 × 4	4	15
	Addition	1920 × 4		
	Linear Activation	1920 × 4		

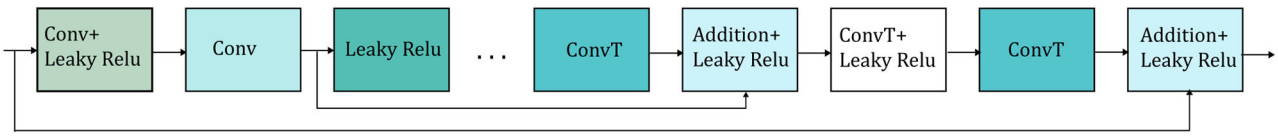

FIGURE 2 | Detailed illustration of the way that the skip connections (represented by the arrows) are applied in the network. Only two skip connections are shown for simplicity. Conv stands for convolution and ConvT for transposed convolution.

The metric was computed for each channel and subsequently averaged over all the ECG channels and test signals.

For real fetal ECG signals there is no ground truth available, because even after the maternal ECG suppression there is still noise present in the signals. Thus, it is impossible to have a gold reference to quantitatively validate the results. Simultaneous scalp recordings may help but they can be performed only during labor. Unfortunately, since our real private dataset was obtained during the second trimester of pregnancy, it was not possible to measure the scalp ECG to have a clean reference. For this dataset, in order to provide some quantitative results along with the qualitative, we decided to generate a surrogate "clean" ground truth signal by calculating the running median of 100 heartbeats. We then measure how well the quality of the denoised signals was enhanced by computing the improvement in SNR performance defined by Equation (2). The metric was calculated for 455 cases, where sufficient QRS complexes were detected for the generation of the "ground truth" signal.

In the Abdominal and Direct Fetal Electrocardiogram Database, since simultaneous scalp measurements are provided together with the non-invasive fetal ECG, the performance of our method was evaluated by comparing with the scalp electrode.

The scalp ECG is however a different lead than the abdominal ones and we cannot compare them directly since, even in case of perfect denoising by our method, the morphology of the ECG will not be the same between different leads. Instead, we estimated a denoised scalp ECG as a linear combination of the four abdominal fetal ECG channels:

$$\hat{X}_{scalp} = a^T X_{denoised}, \;\; a = (X_{denoised} {X_{denoised}}^T)^{-1} X_{denoised} X_{scalp}^T, \tag{3}$$

where X_{scalp} is the [1 x 250] scalp ECG and $\hat{X}_{scalp}$ the [1 x 250] estimation of the scalp ECG from the abdominal fetal ECG channels. The coefficients of the [4 x 1] linear combination, a, were computed on windows of half a second that corresponds to 250 samples. The dimension of $X_{denoised}$ is 4 x 250. Because the scalp ECG measurements contain considerable amount of noise and this could affect the comparison, we denoised the scalp ECG by high pass filtering followed by averaging of 30 ECG complexes. Nevertheless, we provided comparative results both when the estimation was done based on the noisy scalp ECG as well as on the denoised scalp ECG.

Four different quantitative measures were employed for the comparison, the Pearson correlation coefficient (R), the mean squared error (MSE), the mean absolute error (MAE) and the signal-to-noise ratio (SNR). The metrics are defined by the following equations:

$$R = \frac{cov(\hat{X}_{scalp} X_{scalp})}{\sigma_{\hat{X}_{scalp}} \sigma_{X_{scalp}}}, \tag{4}$$

$$MSE = \frac{1}{K} \sum_{i=1}^{K} (X_{scalp_i} - \hat{X}_{scalp_i})^2, \tag{5}$$

$$MAE = \frac{1}{K} \sum_{i=1}^{K} |X_{scalp_i} - \hat{X}_{scalp_i}|, \tag{6}$$

$$SNR = 10 log_{10} \frac{\sum_{i=1}^{K} \left| X_{scalp_i} \right|^2}{\sum_{i=1}^{K} \left| X_{scalp_i} - \hat{X}_{scalp_i} \right|^2}, \tag{7}$$

where cov stands for the covariance, σ the standard deviation and K the length of the signals. The metrics were computed for the five signals of the database and subsequently averaged to obtain one final value.

Reference Methods

Our method was evaluated in comparison with 3 other ECG denoising methods. The first method is the single-channel CNN denoising network, where each fetal ECG channel is denoised separately (21). The second method is a wavelet denoising algorithm that removes the noise by thresholding the detail coefficients after the signal decomposition. The symlet wavelet was selected due to its resemblance with an ECG, while a fixed threshold was used, estimated by the minimax principle (34). The last method is the widely used beat-to-beat averaging method. We selected to average 30 beats similar to the averaging performed by the STAN method (35). The QRS complexes were detected by a Pan Tompkins detector in the clean fetal ECG signals and not the noisy ones because we do not intend to assess the performance of the QRS detector but the performance of the averaging method. However, we should note that it is not guaranteed that the QRS complexes can be accurately estimated in the presence of acute noise.

RESULTS

Performance on Simulated Signals

The improvement in SNR performance of the proposed network in comparison to the other denoising algorithms, for input SNR from −20 to 20 dB, is illustrated in **Figure 3**. As demonstrated in this figure, the CNN network provides a considerable amount of SNR improvement throughout the whole range of input SNR. The proposed method outperforms the beat-to-beat averaging and the wavelet denoising methods for all the input SNR values. This was anticipated because the averaging method does not preserve individual variations among complexes, while our method is capable of doing so. Moreover, the wavelet denoising distorts the signal amplitude, whereas the proposed network preserves it better. The multi-channel network additionally outperforms the single-channel nearly for the whole range of input SNR values. More specifically, for input SNR <0 dB the multichannel algorithm provides an SNR improvement of at least 10 dB with respect to the input signal and at least 2 dB further improvement as compared to the single-channel method. As the input SNR increases the performances of the two methods become gradually comparable, while for input SNR more than 11 dB the single-channel network slightly surpasses the multi-channel. This was something to expect because for signals of lower quality, information from multiple channels will be beneficial for recovering the ECG structure. On the other

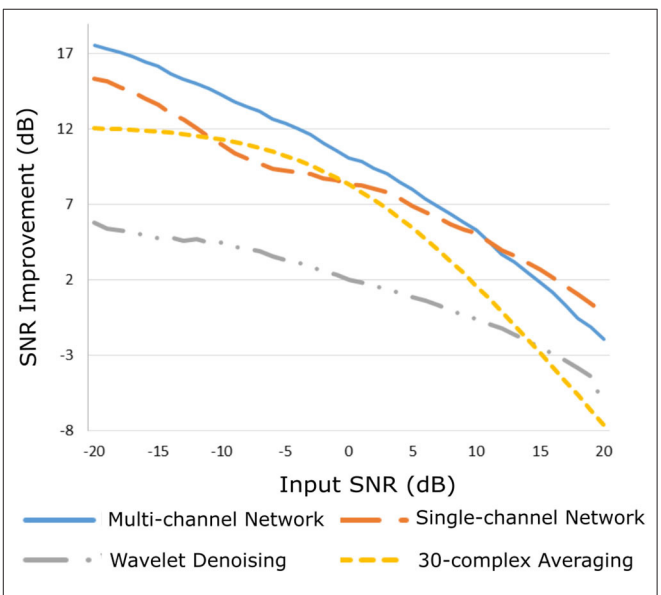

FIGURE 3 | Performance of the proposed multi-channel convolutional network in comparison with other denoising methods in terms of improvement in SNR of the denoised fetal ECG signals when compared with the noisy ones.

hand, if a fetal ECG channel has sufficiently high quality not only the other channels are unnecessary for denoising it but could also slightly affect the quality of the denoised fetal ECG, especially in case their SNR is low. This can be explained better by the following: By using any set of three linearly independent ECG leads, the VCG can be constructed, which is the three-dimensional representation of the electrical activity of the heart. A VCG can explain about roughly 90% of an ECG signal (36). This means that when a signal is reconstructed from different channels, 10% of the signal information should be considered as not reconstructable. In case of very high signal quality, the single-channel denoiser can perform better than the multi-channel since it could theoretically reconstruct 100% of the signal.

By observing **Figure 3**, we see that, for all methods, there is an input SNR for which the denoisers decrease the SNR. This input SNR value is 9, 12, 18, and 20 dB for the wavelet, averaging, multi-channel network and single-channel network denoising methods, respectively. Since it is not common to obtain fetal ECG signals of very high quality (more than 18 dB), we do not consider it as a limitation of our method. We additionally noticed that there is a upgoing trend for the SNR improvement metric as the input SNR decreases. However, we did not test for signals of quality even lower than −20 dB because real fetal ECG signals typically do not have quality less than −20 dB.

Figure 4 depicts two typical denoising results from our test dataset. The SNR values of the signals before and after denoising are provided in **Table 2**. Note that in **Figure 4** the vertical axes limits for the noisy signals differ from those of the ground truth and denoised signals for better visualization. However, the axes limits for the clean and denoised fetal ECG are the same to allow for their comparison. As can be noted, the network suppresses the noise in a great extent for both signals simulated-A (SA) and simulated-B (SB). In the case of signal SA the similarity of the network's output with the clean signals is very high for all channels and all ECG waves are clearly distinguishable. Even for channel 4, with input SNR of −12 dB, the network provides a high-quality result, since it combines all channels to reconstruct it. For signal SB the majority of ECG channels have very low quality (around −9 dB). The SNR after denoising with our network is significantly higher (3.75 dB on average). However, we

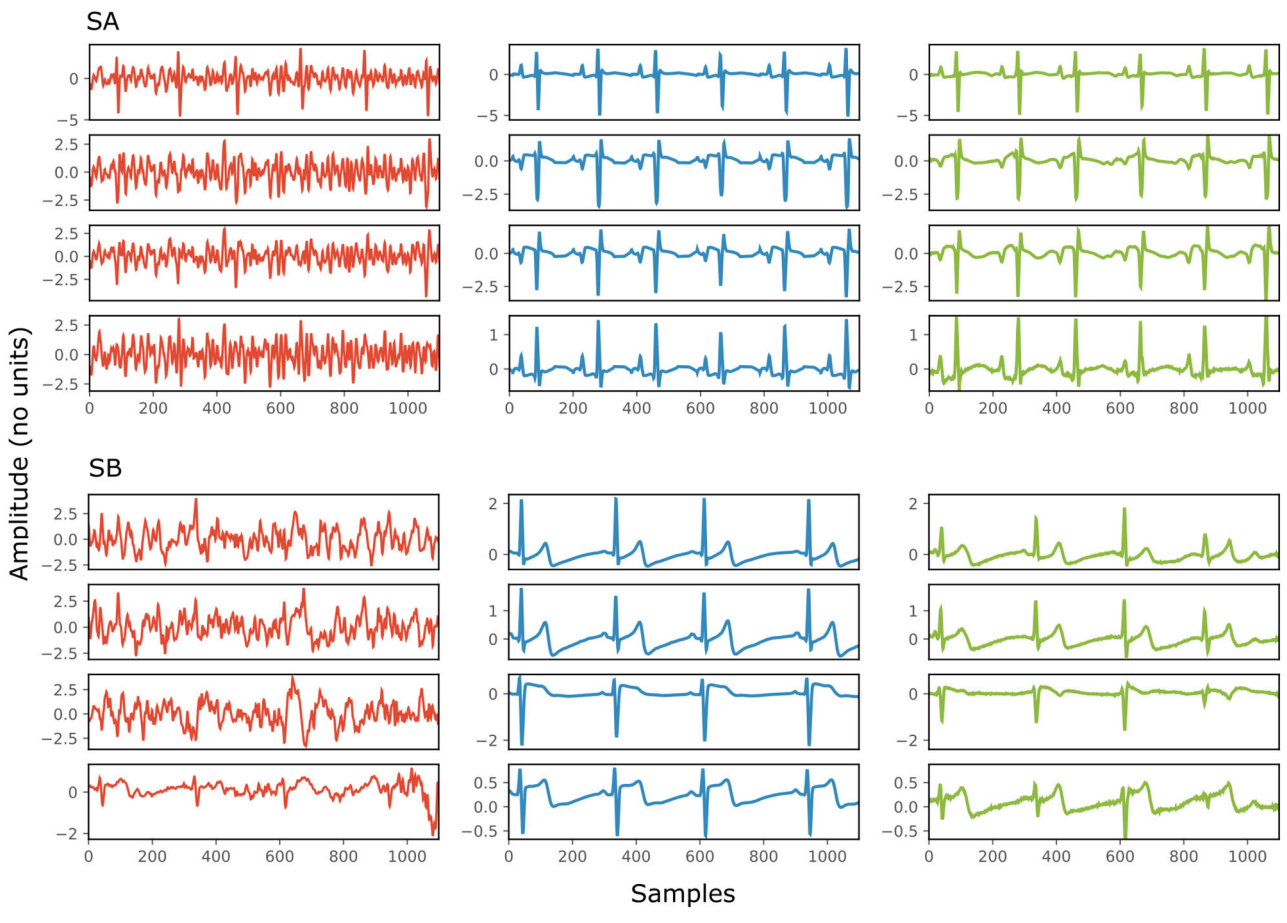

FIGURE 4 | Denoising results by the proposed method for two simulated signals (SA and SB) of the test dataset. For both signals: each panel in the left presents one channel of the noisy four-channel fetal ECG signal (red), in the middle the corresponding channels of the clean signal (blue) are shown and in the right the denoised fetal ECG signal by our network (green). The horizontal axis depicts the samples at 500 Hz, while the vertical the amplitude of the signals. The SNR values of the noisy and the denoised fetal ECG for both signals are given in **Table 2** (SNRin and SNRout, respectively).

TABLE 2 | The SNR values in dB for the four channels of the simulated signals depicted in **Figure 4**, before (SNRin) and after (SNRout) denoising.

Channel	Signal SA		Signal SB	
	SNRin	SNRout	SNRin	SNRout
1	1	17	−9	5
2	−2	13	−9	4
3	−3	14	−10	3
4	−12	8	0	3

notice some distortion on the signal amplitude, while particularly the P-waves are suppressed by the network. Moreover, despite channel 4 having the least amount of noise before entering the network, we observe the least improvement after denoising, evidencing that indeed the network's output is obtained through combination of information from all leads.

Evaluation on Real Fetal ECG Signals

The proposed method was evaluated on our extensive non-invasive fetal ECG dataset (28) and the results are presented in **Figure 5**. **Figure 5** illustrates the improvement in SNR for input SNR ranging from−17 to 1 dB. The input SNR corresponds to the SNR of the noisy fetal ECG signals when we assume that the ground truth signal is the running median of 100 heartbeats. We need to stress that this is not the actual SNR of the signals but merely an approximation of it. In fact, the more noise is present in the signals or the more physiological variation, the less accurate the constructed "clean" signal is. Examining the **Figures 3**, **5**, where the performance in the simulated dataset is illustrated, we observe an analogy between them. In both graphs the multi-channel denoiser surpasses the single-channel for lower input SNR while for higher SNR values the two methods perform comparably. The performance improvement as compared to the single-channel approach is lower for the real signals than for the simulated ones but this might be due to the lack of actual ground truth signals for comparison. By all means the evaluation in this dataset is suboptimal but it provides a performance indicator in a large real dataset.

Figure 6 demonstrates the result of denoising two signals of this database, while **Table 3** provides the corresponding SNR values before and after denoising. Note that the vertical axes limits for the noisy signals differ from the ones of the "clean" and denoised ones for clearer visualization. Both signals in **Figure 6**, especially signal real-B (RB), have a significant amount of noise before denoising (see **Table 3**). The "clean" reference signals as well contain few noise but in most of them the ECG morphology is relatively clear. On the other hand all the possible variations among the successive complexes is lost due to the heavy averaging performed. The multi-channel network achieved a fairly remarkable result in denoising those signals. Comparing the morphology of the denoised with the "clean" reference signals, the various ECG waves and segments correspond relatively well. In this comparison, we acknowledge that the running median of 100 heartbeats is not the gold standard.

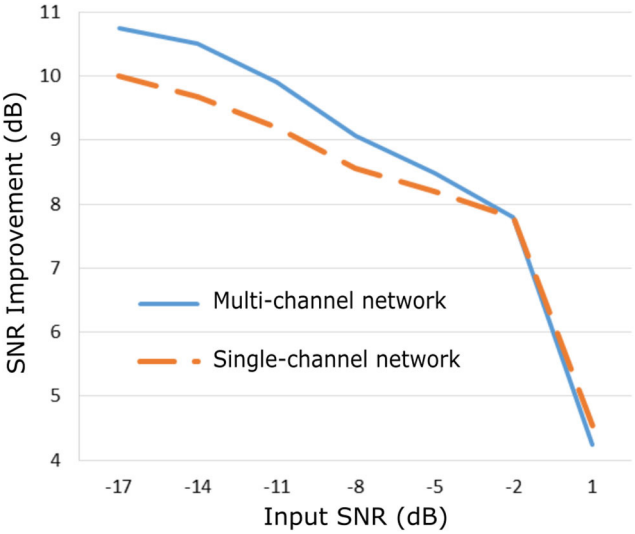

FIGURE 5 | Performance of the proposed fetal ECG denoising method in a large real dataset (28) in terms of improvement in SNR of the denoised signals.

However, the averaging of the heartbeats brings evidence for the location of the ECG waves, especially the P-waves, information that cannot be seen in the noisy signals. It is important to recognize that in the denoised signals by our network, these locations seem to correspond with the locations in the median signals. As a matter of fact, the denoised signals appear to exhibit better quality and clearer morphology than the reference. Some morphological features seem to be distorted, as we can see in **Figure 6** for signal RB. However, the overall performance in those low-quality signals is relatively good.

Figure 7 illustrates the performance for a fetal ECG signal of on non-invasive fetal ECG dataset in comparison to the single-channel network, 30-complex averaging and wavelet denoising. For simplicity we present only one channel out of the four. As shown in the figure, all methods provide a noise-free result. However, our method retains the individual ECG complex differences as opposed to the averaging method and does not distort the signal amplitude as opposed to wavelet denoising. In addition, the morphology of the denoised ECG is clearer in our case. The single-channel network provided a similar result to the multi-channel for this signal.

The performance of the network on the Abdominal and Direct Fetal ECG Database is illustrated in **Table 4**. The scalp ECG was compared with the aforementioned linear combination of abdominal signals, as described in Equation (3). In **Table 4** we provide the results of this comparison for 2 cases; when we used the original scalp ECG and when we denoised it. For each performance metric the values before and after denoising with the multi-channel and single-channel network are presented, while with bold the best performing method is marked.

First, we believe that denoising of the scalp ECG was important to allow for better comparison with the scalp ECG estimation from the denoised abdominal leads. By averaging

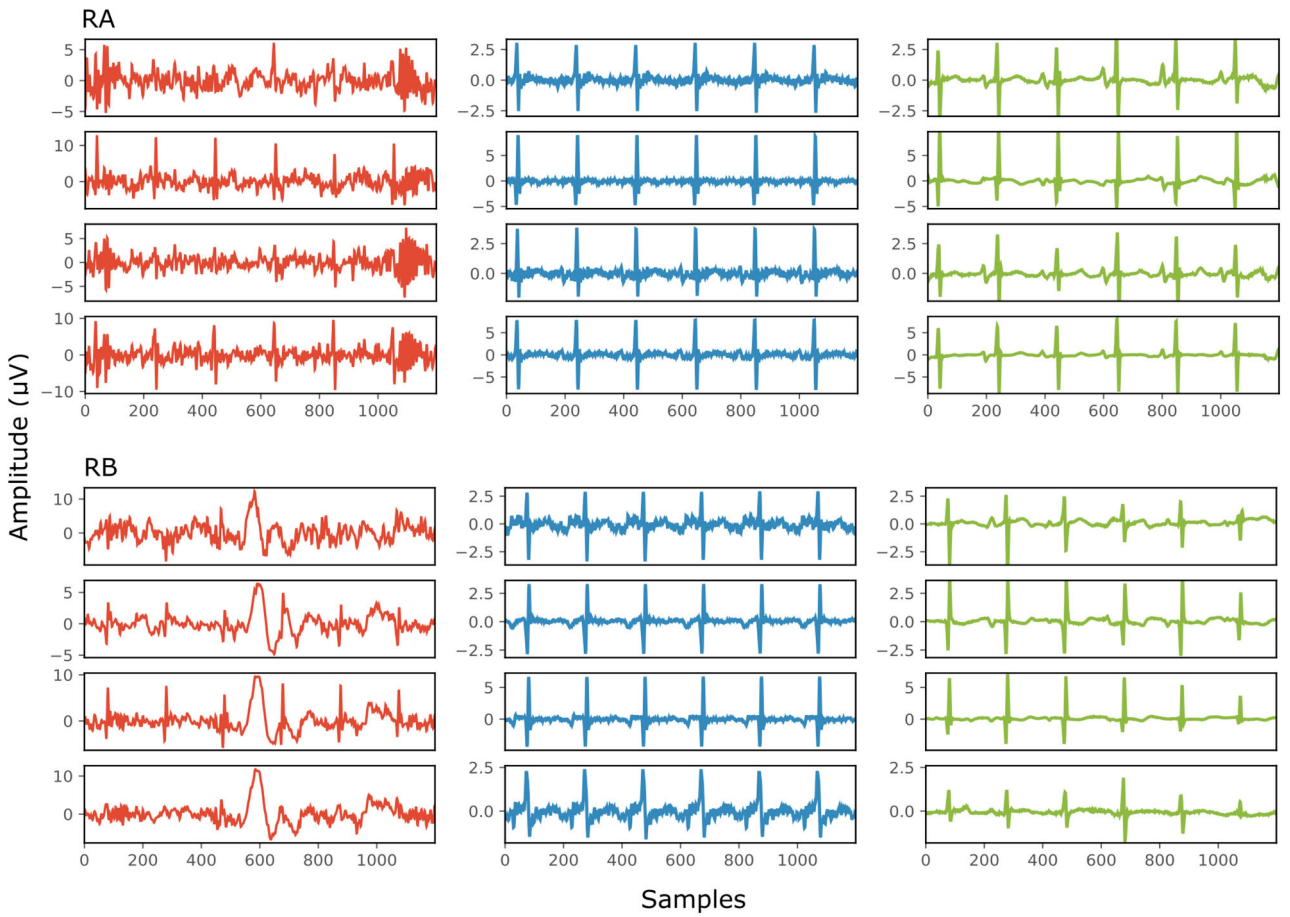

FIGURE 6 | The results of denoising two signals (RA and RB) from our private real fetal ECG dataset (28). For each signal: the noisy four-channel fetal ECG signal extracted from the abdominal measurements is presented in the left (red), the running median of 100 heartbeats for each channel in the middle (blue) and the denoised fetal ECG signal by our network in the right (green). The horizontal axis depicts the samples at 500 Hz, while the vertical the amplitude of the signals in μV. The SNR values of the noisy fetal ECG signals (SNRin) together with the values for the denoised ones (SNRout) are given in **Table 3**.

TABLE 3 | The SNR values in dB for the four channels of the real signals depicted in **Figure 6**, before (SNRin) and after (SNRout) denoising.

Channel	Signal RA		Signal RB	
	SNRin	SNRout	SNRin	SNRout
1	−8	3	−13	3
2	−3	3	−8	5
3	−8	2	−6	6
4	−1	6	−14	1

30 successive ECG complexes we might have lost some morphological variations among the successive beats of the scalp lead but achieved significant quality improvement. Even the scalp ECG approximated by the noisy fetal ECG signals has better resemblance with the denoised scalp lead, e.g., correlation coefficients of 0.74 vs. 0.53. Second, we observe that both the multi-channel and single-channel networks achieve significant quality improvement of the fetal ECG signals for all the metrics presented in **Table 4**. We should note here once more that by no means the scalp estimation is expected to be the same with the scalp ECG even after perfect denoising, because the latter is a different lead than the abdominal leads. Last, the multi-channel network outperforms the single-channel in terms of all computed performance metrics. However, the differences are relatively small. It might be because the extracted fetal ECG signals already have decent quality and, as we have already found in simulated signals, employing multiple channels is more advantageous in cases of signals exhibiting lower SNR. Larger difference was found regarding the MSE metric (62.1 vs. 68.4 μV^2), indicating that the single-channel network may provide more outliers, while the multi-channel a smoother outcome.

Figure 8 provides two qualitative results of the scalp estimation, when fetal ECG denoising is performed with the proposed multi-channel method. In both cases, the scalp estimated by the denoised fetal ECG is free from noise and the individual waves and intervals correspond relatively well to those

of the scalp ECG. We do not expect absolute correspondence, not only because the scalp ECG is a different lead, but also because it was averaged over 30 complexes.

DISCUSSION

We have proposed a CNN network for postprocessing non-invasively extracted multi-channel fetal ECG signals to improve their quality. The non-invasive fetal ECG is substantially contaminated by various noises, even after the application of various signal processing tools proposed in literature, such as maternal ECG suppression. The low quality of the fetal ECG is the principal reason that the applicability of non-invasive fetal electrocardiography in clinical practice is limited. The suggested multi-channel network was trained on a wide dataset of simulated four-channel ECG signals, with SNR ranging from−15 to 15 dB, while it was extensively validated both on simulated as well as on real datasets.

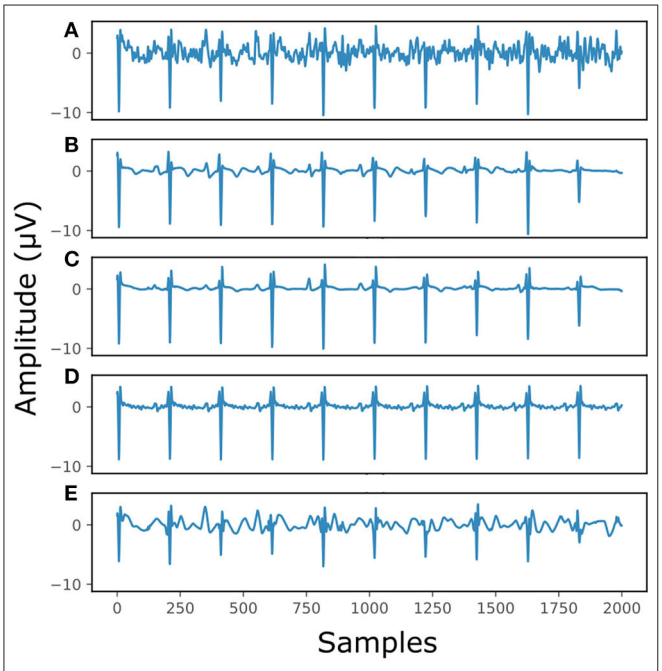

FIGURE 7 | The result of denoising a real fetal ECG signal from our private dataset (28) with different denoising algorithms. For simplicity, only one channel is displayed. The panels show: **(A)** the noisy extracted fetal ECG, **(B)** the denoised signal by the proposed method, **(C)** the denoised signal by the single-channel denoising network, **(D)** the result of 30-complex averaging and **(E)** the result after wavelet denoising. The horizontal axis depicts the samples at 500 Hz, while the vertical the amplitude of the signals in μV.

Experiments on simulated data showed a significant improvement in the quality of the noise-corrupted fetal ECG signals. The network combined information from all the channels to efficiently remove the noise and uncover the ECG signal morphology even in the presence of acute noise. However, the network suppressed some morphological characteristics in cases there was not sufficient content for denoising i.e., when most signal channels were severely corrupted. The multi-channel network outperformed the single-channel (21) in cases of low SNR of the input signals, while for SNR more than 11 dB the single-channel network exhibited slightly better performance. This behavior could be anticipated. A multi-lead signal configuration captures the spatiotemporal nature of the cardiac electrical activity. For low quality signals this is beneficial as more signal information can be exploited to better reconstruct each channel. However, if we wish to denoise a channel that already has high quality, using spatiotemporal content may be not always the best choice. Nevertheless, it is very uncommon in practice to obtain fetal ECG of such high quality. Yet, in case this would happen, the output of the multi-channel network would still be of such quality that it could be used for further clinical interpretations.

The evaluation of our network on a large real fetal ECG dataset showed an analogous behavior to that on the simulated data; for low quality fetal ECG the multi-channel network outperformed the single-channel, while for higher SNR the performances of the networks were comparable. We cannot make a direct comparison because the evaluation method for the real signals was suboptimal. We are aware that the approximation of the ground truth signals with the running average of 100 heartbeats was not very accurate. However, it gave us an indication that the method is efficient in real data too. We additionally presented some qualitative denoising results for two signals of this database in **Figure 6** to support our claim. The network outputted clean denoised signals with good correspondence of the individual

TABLE 4 | Performance of the multi-channel CNN network vs. the single-channel one on the Abdominal and Direct Fetal ECG Database in terms of comparison of the scalp ECG with a scalp estimated from the denoised abdominal fetal ECG.

Metric	Original scalp ECG			Denoised scalp ECG		
	Noisy input	Multi-channel output	Single-channel output	Noisy input	Multi-channel output	Single-channel output
R	0.53	**0.66**	0.65	0.74	**0.87**	0.85
MSE (μV^2)	555.8	**440.7**	449.3	116	**62.1**	68.4
MAE (μV)	15	**12.9**	13	7.3	**5.4**	5.5
SNR (dB)	1.5	**2.7**	2.5	3.7	**6.4**	6.1

The best performing method is marked with bold.

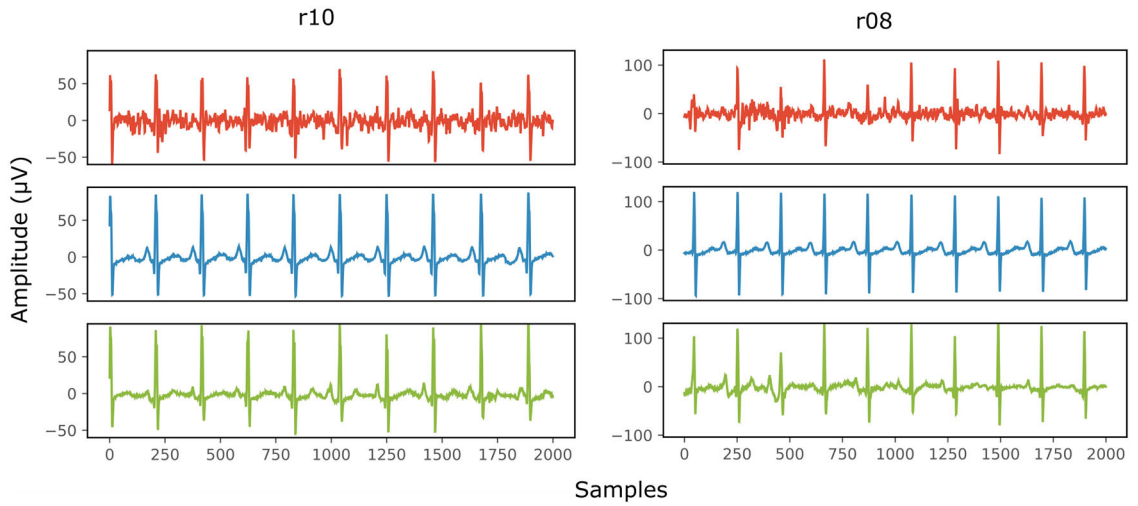

FIGURE 8 | Comparison of the scalp ECG with an estimation of it as a linear combination of the abdominal fetal ECG for 2 records (r10 and r08) of the Abdominal and Direct Fetal ECG Database. In the first row (red) the scalp ECG estimated from the noisy fetal ECG is presented, in the second (blue) the ECG as measured by the scalp electrode (running average of 30 complexes) and in the last one (green) the scalp ECG estimated from the denoised fetal ECG by the proposed network. The horizontal axis depicts the samples at 500 Hz, while the vertical the amplitude of the signals in μV.

ECG waves between the reference and denoised signals. A few recordings in our dataset had input SNR that was even lower than−17 dB. Based on visual analysis of the output of our proposed denoiser, we could argue that the performance of the denoiser breaks down at these very low signal quality levels and the network is no longer capable of reconstructing a reliable fetal ECG. This limitation probably comes from the fact that the network was trained for input SNR range of−15 to 15 dB. Thus, the network did not learn to remove efficiently the noise when the quality of the signals is even lower. This indicates that we might need to perform experiments for a wider SNR input range. However, the capacity of the network might no longer be sufficient for handling such an ample range of signal qualities and further research is needed to evaluate this.

The CNN network was additionally evaluated in the Abdominal and Direct Fetal ECG Database. Simultaneously recorded scalp ECGs were compared to an estimated scalp ECG from the denoised abdominal channels, also here demonstrating that the method can provide significant quality improvement of the noisy fetal ECG signals. Comparison of the performances of the multi-channel and single-channel networks for this database, revealed that they achieve comparable results, probably because the input signals were of relatively good quality. It is difficult to compare the performances between the two real datasets for several reasons. Most importantly, the sizes of the two datasets differ a lot (455 vs. 5) and so do the gestational ages of the subjects (18-24 vs. 38-41 weeks).

As mentioned in the introduction and also in (21), the shortcoming of denoising single-channel fetal ECG with a convolutional network is that the network can output signals that look as if they were ideally denoised, but that can have "fake" waves that can differ both in location and polarity when compared to the actual ECG waves. This happens mostly when the quality of the input signals is relatively low and the network, not having enough signal information, reconstructs a clean signal from unreliable information in the encoded latent space. We demonstrated that by employing multichannel signals this problem is eliminated to a large extent. When the quality of the signals is very low, the amplitude of the small signal waves, like the P-wave and T-wave, and less often of the R-peaks in the denoised signals can be distorted rather than "fake." This means that some waves may be virtually absent, or the output does not even resemble an ECG anymore. This makes the method safer to use in clinical practice, because clinicians will typically discard a distorted signal but a signal that looks like a high-quality ECG but in fact contains "fake" information might lead to erroneous decision-making.

To summarize, we have shown the potential of deep CNNs for removing noise from non-invasive multi-lead fetal ECG. We validated the method on a wide dataset of simulated but also real recordings with both early as well as late gestational ages (18 to 24 and 38 to 41 weeks). Primarily, we demonstrated that employing multi-channel information for denoising does not only lead to more clean signals but also to more reliable results, when compared to single-channel information. The main advantage of the method is that, as opposed to the widely used averaging method, no prior processing of the signal is needed to extract the locations of the R-peaks and variations in ECG morphology among consecutive heartbeats are preserved. This is especially important in case that arrhythmias are present. Up to now, arrhythmia is assessed through echocardiography because the averaging that was performed to enhance the quality of the fetal ECG hinders its application for arrhythmia analysis. Moreover, the quality of the denoised signals is high enough to allow for measuring the timing of intervals, like the PR and QT interval. However, in order to confirm this, we need to perform a

thorough comparison of the ECG intervals between the denoised and the clean signals. If we wish to obtain reliable results, a large annotated dataset is necessary, but this requires time and experts to perform these annotations.

Certainly, there is room for improvement of our method. Most importantly, the capacity of the network could be further increased to handle even more noisy signals. Moreover, we can explore denoising directly the raw abdominal signals, without cancelling the maternal ECG. Most probably a more complex network architecture is needed for such a task and appropriate data for training.

CONCLUSION

An end-to-end trained deep CNN network was presented for denoising of fetal ECG signals. Convolutions and transposed convolutions were combined in the network, modeling the denoising problem as an encoding of primary signal content and subsequent decoding to recover details. Essentially, we proposed to employ spatiotemporal information in the ECG signal by using multiple ECG leads simultaneously as input to the network. The network then learned how to combine the input channels and deliver a reliable clean ECG as output. Experiments on simulated as well as in real data showed that the network can achieve a substantial quality improvement of the noisy signals and outperform a single-channel alternative.

DATA AVAILABILITY STATEMENT

Data are available from the Data Governance Board of the Máxima Medical Centre to researchers who can demonstrate that they are qualified to use confidential data. Requests to access the data should be addressed to the corresponding author or the Data Governance Board at the Máxima Medical Centre (J. Luime, j.luime@mmc.nl). Other data used in the study can be found in: https://archive.physionet.org/physiobank/database/adfecgdb/?C=S;O=A.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The study was approved by the Máxima Medical Centre institutional review board (NL48535.015.14). Participants were included in the study after written informed consent had been obtained. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EF is a Ph.D. student and has done the work of the paper, while RV closely supervised the work while giving ideas, feedback, and guidance.

REFERENCES

1. van Laar JO, Warmerdam GJ, Verdumen KM, Vullings R, Peters CH, Houterman S, et al. Fetal heart rate variability during pregnancy, obtained from non-invasive electrocardiogram recordings. *Acta Obstet Gynecol Scand.* (2014) 93:93–101. doi: 10.1111/aogs.12286
2. Amer-Wåhlin I, Hellsten C, Norén H, Hagberg H, Herbst A, Kjellmer I. Cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram for intrapartum fetal monitoring: a Swedish randomised controlled trial. *Lancet.* (2001) 35:8534–8. doi: 10.1016/s0140-6736(01)05703-8
3. Behar J, Andreotti F, Zaunseder S, Oster J, Clifford GD. A practical guide to non-invasive foetal electrocardiogram extraction and analysis. *Physiol Meas.* (2016) 37R1-R35. doi: 10.1088/0967-3334/37/5/R1
4. Sameni R, Clifford GD. A review of fetal ECG signal processing; issues and promising directions. *Open Pacing Electrophysiol Ther.* (2010) 3:4–20. doi: 10.2174/1876536X01003010004
5. Lindecrantz K. *Processing of the Fetal ECG; An Implementation of a Dedicated Real Time Microprocessor System*. Chalmers University of Technology, Gothenburg. (1983)
6. Ahmadi M, Ayat M, Assaleh K, Al-Nashash H. Fetal ECG signal enhancement using polynomial classifiers and wavelet denoising. In: *2008 Cairo International Biomedical Engineering Conference.* Cairo (2008). doi: 10.1109/CIBEC.2008.4786095
7. Baldazzi G, Sulas E, Brungiu E, Urru M, Tumbarello R, Raffo L, et al. Wavelet-based post-processing methods for the enhancement of non-invasive fetal ECG. In: *2019 Computing in Cardiology*, Singapore (2019) doi: 10.23919/CinC49843.2019.9005921
8. Fotiadou E, van Laar JOEH, Oei SG, Vullings R. Enhancement of low-quality fetal electrocardiogram based on time-sequenced adaptive filtering. *Med Biol Eng Comput.* (2018) 56:231–323. doi: 10.1007/s11517-018-1862-8
9. Mao X, Shen C, Yang Y. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In: *NIPS*. Barcelona (2016).
10. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. *Trans Image Process.* (2017) 26:3142–55. doi: 10.1109/TIP.2017.2662206
11. Jain V, Seung S. Natural image denoising with convolutional networks. In: *Advances in Neural Information Processing Systems*. Vancouver, BC (2008).
12. Maas A, Le QV, O'Neil TM, Vinyals O, Nguyen P, Ng AY. Recurrent neural networks for noise reduction in robust. In: *Interspeech.* (2012).
13. Lu X, Tsao Y, Matsuda S, Hor C. Speech enhancement based on deep denoising autoencoder. In:*Interspeech*. (2013).
14. Antczak K. Deep recurrent neural networks for ECG signal denoising. *arxiv [Preprint] arXiv:1807.11551.*. (2018)
15. Xiong P, Wang H, Liu M, Zhou S, Hou Z, Liu X. ECG signal enhancement based on improved denoising auto-encoder. *Eng Appl Artif Intell.* (2016) 52:194–202. doi: 10.1016/j.engappai.2016.02.015
16. Zhong W, Liao L, Guo X, Wang G. A deep learning approach for fetal QRS complex detection. *Physiol Meas.* (2018) 39:045004. doi: 10.1088/1361-6579/aab297
17. Lee JS, Seo M, Kim SW, Choi M. Fetal QRS detection based on convolutional neural networks in noninvasive fetal electrocardiogram. In: *4th International Conference on Frontiers of Signal Processing*. Poitiers (2018). doi: 10.1109/ICFSP.2018.8552074
18. Muduli PR, Gunukula RR, Mukherjee A. A deep learning approach to fetal-ECG signal reconstruction. In: *2016 Twenty Second National Conference on Communication (NCC),* Guwahati (2016). doi: 10.1109/NCC.2016.7561206
19. Zhong W, Guo X, Wang G. Non-invasive fetal electrocardiography denoising using deep convolutional encoder-decoder networks. In: *Proceedings of 2019 Chinese Intelligent Systems Conference*. Singapore (2019). doi: 10.1007/978-981-32-9682-4_1
20. Fotiadou E, Konopczyński T, Hesser J, Vullings R. Deep convolutional encoder-decoder framework for fetal ECG signal

denoising. In: *2019 Computing in Cardiology (CinC)*. Singapore (2019). doi: 10.23919/CinC49843.2019.9005722
21. Fotiadou E, Konopczyński T, Hesser J, Vullings R. End-to-end trained encoder–decoder convolutional neural network for fetal electrocardiogram signal denoising. *Physiol Meas.* (2020) 41:15005. doi: 10.1088/1361-6579/ab69b9
22. Behar J, Andreotti F, Zaunseder S, Li Q, Oster J, Clifford GD. An ECG simulator for generating maternal-foetal activity mixtures on abdominal ECG recordings. *Physiol Meas.* (2014) 35:1537–50. doi: 10.1088/0967-3334/35/8/1537
23. Goldberger L, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mar RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation.* (2000) 101:e215–20. doi: 10.1161/01.cir.101.23.e215
24. McSharry PE, Clifford GD, Tarassenko L, Smith L. A dynamical model for generating synthetic electrocardiogram signals. *IEEE Trans Biomed Eng.* (2003) 50:289–94. doi: 10.1109/TBME.2003.808805
25. Sameni R, Clifford GD, Jutten C, Shamsollahi BM. Multichannel ECG noise modeling: application to maternal fetal ECG signals. *EURASIP J Adv Sig Pr.* (2007) 2007:43407. doi: 10.1155/2007/43407
26. Andreotti F, Behar J, Zaunseder S, Oster J, Clifford GD. An open-source framework for stress-testing non-invasive foetal ECG extraction algorithms. *Physiol Meas.* (2016) 37:627–48. doi: 10.1088/0967-3334/37/5/627
27. Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik.* (1995) 40:s1. doi: 10.1515/bmte.1995.40.s1.317
28. Verdurmen KMJ, Lempersz C, Vullings R, Schroer C. Normal ranges for fetal electrocardiogram values for the healthy fetus of 18–24 weeks of gestation: a prospective cohort study. *BMC Preg Childbirth.* (2016) 16:227. doi: 10.1186/s12884-016-1021-x
29. Lempersz C, van Laar JO, Clur AB, Verdurmen KM, Warmerdam GJ, van der Post J, et al. The standardized 12-lead fetal electrocardiogram of the healthy fetus in mid-pregnancy: a cross-sectional study. *PLoS ONE.* (2020) 15:e0232606. doi: 10.1371/journal.pone.0232606
30. Jezewski J, Matonia A, Kupka T, Roj D, Czabanski R. Determination of fetal heart rate from abdominal signals: Evaluation of beat-to-beat accuracy in relation to the direct fetal electrocardiogram. *Biomed Engin.* (2012) 57:383–94. doi: 10.1515/bmt-2011-0130
31. Varanini M, Tartarisco G, Billeci L, Macerata A, Pioggia G, Balocchi R. A multi-step approach for non-invasive fetal ECG analysis. In: *Computing in Cardiology 2013*. Zaragosa (2013).
32. Burger HC, Schuler CJ, Harmeling S. Image denoising: Can plain neural networks compete with BM3D? In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Providence, RI (2012) doi: 10.1109/CVPR.2012.6247952
33. Kingma D, Ba J. Adam: A method for stochastic optimization. In: *International Conference on Learning Representations*. San Diego, CA (2015).
34. Sardy S. Minimax threshold for denoising complex signals with Waveshrink. *IEEE Trans Signal Process.* (2000) 48:1023–8. doi: 10.1109/78.827536
35. Hulsenboom DJ, Verdurmen KM, Vullings R, van der Hout-van der Jagt MB, Kwee A, Joe van Laar H. Relative versus absolute rises in T/QRS ratio by ST analysis of fetal electrocardiograms in labour: A case-control pilot study. *PLoS ONE.* (2019) 14:0124357. doi: 10.1371/journal.pone.0214357
36. Geselowitz DB. On the theory of the electrocardiogram. *Proc IEEE.* (1989) 77:857–76. doi: 10.1109/5.29327

Conflict of Interest: RV has shares in Nemo Healthcare BV, The Netherlands.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

OPINION
published: 27 August 2020
doi: 10.3389/fped.2020.00503

Electronic Fetal Monitoring–Prevention or Rescue?

*Barry S. Schifrin**

Department of Obstetrics & Gynecology, Western University of Health Sciences, Pomona, CA, United States

Keywords: EFM, classification of FHR patterns, fetal acidemia, obstetrical malpractice, defensive medicine, malpractice awards

INTRODUCTION

This commentary represents a response to two recent contributions to the literature on electronic fetal monitoring (EFM) (1, 2). One article by Hirsch, raises concerns about the value of fetal monitoring in light of a very large judicial award of $50 million against an obstetrical service for a "brain-damaged baby" (1). In the other article entitled: "Fetal Heart Rate Monitoring: Still a Mystery More than Half a Century Later," the authors present a Category II EFM tracing which creates for them uncertainty about whether the "fetal acid-base balance may be affected" and whether they may await spontaneous vaginal delivery or must consider more expedient operative delivery (2).

BEWILDERING FHR PATTERNS

The authors of the two articles share a bewilderment about the meaning of fetal heart rate (FHR) patterns and the problems of responding responsibly to them. They observe that despite the longevity of its implementation, the ubiquity of its use, and modifications to the classification of FHR patterns over the past 50 years, the interpretation of FHR patterns has continued to "befuddle obstetric care providers."

In these responses, the authors are reflecting a broader bewilderment in the society of obstetricians. In the most recent coverage of the topic in Up-To-Date the author finds no unequivocal benefit to the use of EFM; further insisting that it is equivalent to intermittent auscultation (3). In a publication on the evaluation and response to Category II patterns, eighteen well-known authors confess that "As a medical community, we seem to know less than we thought we did 30 years ago regarding the utility of this ubiquitous technology." They also aver that "Unfortunately, this body of work [EFM research] has primarily served to raise more questions than it has answered" (4). In a subsequent study their proposed scheme to manage Category II patterns was found to be of very limited benefit (5). One is reminded of Churchill's description of Russia as a "riddle wrapped in a mystery, inside an enigma."

The "befuddlement" of obstetric care providers comes not only from the promulgation of what may be considered "defensive" classifications of FHR patterns (including—Categories I–III from ACOG), but also from a monolithic, and ill-considered view of the role of acidemia in the provenance of adverse fetal neurological outcome.

THE CLASSIFICATION OF FHR PATTERNS

The currently popular, three-category classification of fetal heart rate patterns in the United States (Category I–III) was introduced in 2008 without proper vetting and without attention to fundamental physiological principles (6). These comments may be applied to classifications from other national/regional organizations as well. These classifications, tied exclusively to the estimate of fetal acidemia and ensuing hypoxic-ischemic injury, attributes little importance to the proper

OPEN ACCESS

Edited by:
Ines Maria Nunes, North Maternal Child Center Albino Aroso (CMIN), Portugal

Reviewed by:
Edwin Chandraharan, Global Academy of Medical Education and Training, United Kingdom
Susanna Timonen, Turku University Hospital, Finland
Austin Ugwumadu, University of London, United Kingdom
Sabaratnam Arulkumaran, University of London, United Kingdom

***Correspondence:**
Barry S. Schifrin
bschifrinmd@aol.com

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *06 April 2020*
Accepted: *16 July 2020*
Published: *27 August 2020*

Citation:
Schifrin BS (2020) Electronic Fetal Monitoring–Prevention or Rescue? Front. Pediatr. 8:503. doi: 10.3389/fped.2020.00503

assessment of uterine activity and ignores the concept of using the individual FHR pattern, over time, as its own control. Moreover, it imposes arbitrary definitions of tachycardia and bradycardia and takes no instruction from the recovery of the fetus from the individual deceleration or its evolution over time; there is no recognition of the importance of fetal behavior or the potential for the prospective identification of fetal neurological injury or intracranial hemorrhage (7, 8). They do not obey physiological principles (9).

Category II patterns offer disparate combinations of (1) decelerations (early, late or variable) with normal baseline features or (2) abnormal baseline features (altered variability, absent accelerations, tachycardia, bradycardia) without decelerations. Given the breadth of physiological and pathological conditions that may present with a Category II tracing or equivalent designation in other classifications (cord compression, head compression, placental insufficiency, medication effects, prematurity, fetal sleep cycles, existing injury, anomaly, etc.), it is unreasonable to consider that the metabolic status or the tissue oxygen reserve of each fetus, or the time to decompensation is the same. Indeed, "Category II" pattern may reflect a normal healthy fetus, but it does not exclude either fetal acidosis or fetal neurological injury (10–13).

Combining these disparate features and etiologies into a single classification and offering vague guidelines for their management including "continued surveillance and reevaluation" (3, 14) appears to have contributed to the "conundrum" for those providers trying to decide what the ubiquitous (identified in 70–80% of fetuses during labor) (15) Category II tracings mean, how to respond to them, it and how to counsel patients.

IMPLEMENTATION OF THE CATEGORY SYSTEM

The argument that the introduction of EFM was not accompanied by rigorous studies can also be made about the adoption of the various systems of classifying FHR patterns and tying it exclusively to the presence or absence of fetal acidemia. A greater deficiency was the failure to understand the provenance of intrapartum fetal injury based on the assessment of both immediate and long-term outcome, not just injury associated with a very low pH. Although ACOG guidelines, at least, accept the evolution of Category I to Category III as confirmation of an intrapartum injury, most babies injured during labor do not have severely pathological patterns (11), nor are they acidemic at birth. The de rigueur requirement of acidosis to make the correlation between fetal heart rate patterns and injury made a virtue out of necessity in that no measurements of greater relevance, such as fetal blood pressure or cerebral perfusion, were available.

Why is EFM not beneficial? It is universally agreed that fetal heart rate patterns reliably detect fetal hypoxia and are strongly related to adverse outcomes (16, 17). If the test of its value, however, rests with the correlation with pH or base deficit (BD) at the time of birth and not with long-term outcome, then the wrong question is being asked. On the other hand, if EFM has no preventive value, except to increase the cesarean section rate, what can possibly be the justification for either it or intermittent auscultation?

One can only agree with the enlightened notion of Andrews and Tivo that "measures should be employed in an effort to convert the category II to a category I tracing" (2). In this recommendation lies the likely reengineering of the approach to EFM as an instrument of preventive care rather than one geared to rescuing the fetus from an obviously hostile, presumably acidemic, environment (18). In this respect, it is necessary to place monitoring in its proper context by a careful evaluation of maternal, fetal and obstetrical risk factors. It is especially necessary to ensure adequate fetal reserve at the outset of monitoring; to scrupulously avoid excessive uterine activity (an independent risk factor for adverse outcome) irrespective of heart rate pattern and to titrate the mother's expulsive efforts according to the response of the fetus. Expulsive efforts dramatically increase further the intrauterine pressure and the distortion (molding) of the fetal head with potential compromise of fetal cerebral blood flow. These should be considered as primary instruments to prevent or improve abnormal FHR patterns and minimize the need for urgent intervention. Further, these initiatives must be taken as early as possible, and assessed for trajectory with each contraction. Failure to observe improvement after a reasonable number of contractions, not time, gives credence to the notion of a fetus on a trajectory of decreasing fetal reserve, irrespective of whether some specific pH or base BD value has been reached. There appears to be no clinical virtue to seeing how close one comes to catastrophe before intervening (rescuing). Indeed, withholding intervention until the pattern reaches Category III makes the determination of fetal acidemia more important than a normal fetal outcome.

MALPRACTICE ALLEGATIONS

Finally, we come to the proverbial elephant in the room—the allegation of obstetrical malpractice that rests with the interpretation of the EFM tracing—a concern so ubiquitous that it appears in many if not most articles on fetal monitoring. This is not without cause; the substandard response to FHR patterns is a conspicuous mainstay of preventable injury worldwide whether the tribunal is the courtroom or organizational review (19, 20).

In his article, Hirsch provides neither the tracing nor sufficient information to evaluate its role in the outcome of the patient or the medico-legal encounter. We are also not informed of the positions of the attorneys for either side or most critically, the credibility of the various witnesses including the experts and the defendant. One conclusion that can be drawn without knowing any of these details, however, is that the jury was dismayed by the deportment of the defense.

The $50 million award that "hinged on the disputed interpretation of the fetal heart rate pattern," was certainly an unusually large award and well above the "average" $1 million estimated average lifetime costs of health care, including costs for productivity, and for social outlays (Morbidity and Mortality Weekly Report-MMWR). These estimates do not include out of

pocket expenses, lost wages, emergency room visits, over-the-counter medications, caregiving expenses, among others which are greater than average if the infant with severe brain-damage has a close to normal life expectancy. Nor does it include the toll paid by the participants on both sides of the often lengthy, often dispiriting, process of discovery that characterizes Western litigation. Nevertheless, the questions must be asked: what fact pattern could have so impelled the jury not only to a verdict of negligence, but to such a very large award? And does this not represent a breakdown of the legal system—a lottery? Was it the high-flown oratory of the plaintiff's expert? Andrews and Tivo allege that "only individuals with *post-hoc* knowledge of the neonatal outcome (plaintiff experts?) seem to be proficient at interpretation" (2). Was it the severity of the injury? Was the afflicted child brought into court to prey on the sympathies of the jurors? My own experience suggests that it is not negligence alone that accounts for such rare, "runaway" awards, but deception by the defense; by distorting the evidence about the value of both EFM and hands-on obstetrical care as vouchsafed by "authoritative" sources. The jury has no other mechanism for penalizing the defense team—it has no option to say: "We award the plaintiff a fraction of the amount that we were considering and perhaps he/she will need, but we also want to show how disappointed we were by the actions of the defense that diminished our notion of the honor of the medical profession."

Consider the following heart-felt testimony of an empathetic, defendant physician: "I am so sorry about this outcome; I grieve for the child and for the parents. I have never intended to harm anyone, much less a patient. I believed, then and now, that I was acting reasonably under the circumstances. Both for the patient's sake and my own, I wish we had the day to do over again." Imagine the response of the very human jurors to this confession, not of medical error, but of human value, even of fallibility. Even in the dock (the witness stand) the doctor was trying to heal. Sometimes substandard care harms babies and mothers and is deserving of honest reckoning and adequate compensation. An award, however generous, can perhaps palliate the injury or perhaps lessen the heartache, but it is the humanity of the defendant that not only prevents "runaway" verdicts, but most importantly, also offers some solace (healing) to the parents while holding open the option of learning something from the experience that will benefit a future patient.

CONCLUSION

Obstetrical health care providers continue to look for guidance in the poorly conceived classifications of FHR patterns largely unrelated to our understanding of fetal-maternal physiology and predicated on the notion of EFM as an instrument of rescue from "threatening" acidemia. We should acknowledge that these constructs and the "vagueness" of the management guidelines better protect the physician from the allegation of malpractice than the fetus from the potentially harmful stresses of labor and delivery. As the emotions present in these two contemporaneous articles clearly convey, we are all paying a dear price for that approach.

The quotation from Churchill cited above ends with: "but perhaps there is a key." The "key" is to have a better understanding of the language (physiology) and trends of fetal heart rate patterns and to broadly adopt a less defensive posture that reorients our priorities so that we are more offended by bad outcomes than the specter of malpractice litigation. We must increase our support for parents before, during and after pregnancy and embrace the notion that what we do as obstetrical care providers does matter, perhaps long after the pregnancy is over. Health care providers also need support, but the classification of FHR patterns cannot immunize us against accountability or empathy.

AUTHOR CONTRIBUTIONS

BS conceived, wrote, and approved the final version of the manuscript.

REFERENCES

1. Hirsch E. Electronic fetal monitoring to prevent fetal brain injury: a ubiquitous yet flawed tool. *JAMA*. (2019) 322:611–2. doi: 10.1001/jama.2019.8918
2. Andrews WW, Tita ATN. Fetal heart rate monitoring: still a mystery more than half a century later. *Obstet Gynecol.* (2020) 135:469–71. doi: 10.1097/AOG.0000000000003674
3. Miller D. *Overview of Fetal Heart Rate Assessment: Up to Date*. Barrs (2020). Available online at: https://www.uptodate.com/contents/overview-of-intrapartum-fetal-heart-rate-assessment
4. Clark SL, Nageotte MP, Garite TJ, Freeman RK, Miller DA, Simpson KR, et al. Intrapartum management of category II fetal heart rate tracings: towards standardization of care. *Am J Obstet Gynecol.* (2013) 209:89–97. doi: 10.1016/j.ajog.2013.04.030
5. 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) 216:163 e1–6. doi: 10.1016/j.ajog.2016.10.009
6. Macones GA, Hankins GD, Spong CY, Hauth J, Moore T. The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. *J Obstet Gynecol Neonatal Nurs.* (2008) 37:510–5. doi: 10.1111/j.1552-6909.2008.00284.x
7. Schifrin BS, Ater S. Fetal hypoxic and ischemic injuries. *Curr Opin Obstet Gynecol.* (2006) 18:112–22. doi: 10.1097/01.gco.0000192984.15095.7c
8. Vintzileos AM, Smulian JC. Decelerations, tachycardia, and decreased variability: have we overlooked the significance of longitudinal fetal heart rate changes for detecting intrapartum fetal hypoxia? *Am J Obstet Gynecol.* (2016) 215:261–4. doi: 10.1016/j.ajog.2016.05.046
9. Jia YJ, Chen X, Cui HY, Whelehan V, Archer A, Chandraharan E. Physiological CTG interpretation: the significance of baseline fetal heart rate changes after the onset of decelerations and associated perinatal outcomes. *J Matern Fetal Neonatal Med.* (2019) 12:1–6. doi: 10.1080/14767058.2019.1666819
10. Cahill AG, Mathur AM, Smyser CD, McKinstry RC, Roehl KA, Lopez JD, et al. Neurologic injury in acidemic term infants. *Am J Perinatol.* (2017) 34:668–75. doi: 10.1055/s-0036-1597135
11. Frey HA, Liu X, Lynch CD, Musindi W, Samuels P, Rood KM, et al. An evaluation of fetal heart rate characteristics associated with

neonatal encephalopathy: a case-control study. *BJOG.* (2018) 125:1480–7. doi: 10.1111/1471-0528.15222
12. Evans MI, Britt DW, Eden RD, Gallagher P, Evans SM, Schifrin BS. The fetal reserve index significantly outperforms ACOG category system in predicting cord blood base excess and pH: a methodological failure of the category system. *Reprod Sci.* (2019) 26:858–63. doi: 10.1177/1933719119833796
13. Evans MI, Eden RD, Britt DW, Evans SM, Schifrin BS. Re-engineering the interpretation of electronic fetal monitoring to identify reversible risk for cerebral palsy: a case control series. *J Matern Fetal Neonatal Med.* (2019) 32:2561–9. doi: 10.1080/14767058.2018.1441283
14. American College of Obstetricians and Gynecologists. Practice Bulletin No. 116: Management of intrapartum fetal heart rate tracings. *Obstet Gynecol.* (2010) 116:1232–40. doi: 10.1097/AOG.0b013e3182004fa9
15. Jackson M, Holmgren CM, Esplin MS, Henry E, Varner MW. Frequency of fetal heart rate categories and short-term neonatal outcome. *Obstet Gynecol.* (2011) 118:803–8. doi: 10.1097/AOG.0b013e31822f1b50
16. Spong CY. Electronic fetal heart rate monitoring: another look. *Obstet Gynecol.* (2008) 112:506–7. doi: 10.1097/AOG.0b013e318185f872
17. Williams RL, Hawes WE. Cesarean section, fetal monitoring, and perinatal mortality in California. *Am J Public Health.* (1979) 69:864–70. doi: 10.2105/AJPH.69.9.864
18. Eden RD, Evans MI, Evans SM, Schifrin BS. Reengineering electronic fetal monitoring interpretation: using the fetal reserve index to anticipate the need for emergent operative delivery. *Reprod Sci.* (2018) 25:487–97. doi: 10.1177/1933719117737849
19. Berglund S, Grunewald C, Pettersson H, Cnattingius S. Severe asphyxia due to delivery-related malpractice in Sweden 1990–2005. *BJOG.* (2008) 115:316–23. doi: 10.1111/j.1471-0528.2007.01602.x
20. Buchmann EJ, Velaphi SC. Confidential enquiries into hypoxic ischaemic encephalopathy. *Best Pract Res Clin Obstet Gynaecol.* (2009) 23:357–68. doi: 10.1016/j.bpobgyn.2008.12.004

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

SYSTEMATIC REVIEW
published: 11 September 2020
doi: 10.3389/fped.2020.00584

Monitoring Fetal Electroencephalogram Intrapartum: A Systematic Literature Review

*Aude Castel[1], Yael S. Frank[2], John Feltner[3], Floyd B. Karp[4], Catherine M. Albright[2] and Martin G. Frasch[2,5]**

[1] Department of Clinical Sciences, Faculty of Veterinary Medicine, Université de Montréal, Montreal, QC, Canada, [2] Department of Obstetrics and Gynecology, University of Washington, Seattle, WA, United States, [3] Department of Pediatrics, University of Washington, Seattle, WA, United States, [4] School of Pharmacy, University of Washington, Seattle, WA, United States, [5] Center on Human Development and Disability, University of Washington, Seattle, WA, United States

Background: Studies about the feasibility of monitoring fetal electroencephalogram (fEEG) during labor began in the early 1940s. By the 1970s, clear diagnostic and prognostic benefits from intrapartum fEEG monitoring were reported, but until today, this monitoring technology has remained a curiosity.

Objectives: Our goal was to review the studies reporting the use of fEEG including the insights from interpreting fEEG patterns in response to uterine contractions during labor. We also used the most relevant information gathered from clinical studies to provide recommendations for enrollment in the unique environment of a labor and delivery unit.

Data Sources: PubMed.

Eligibility Criteria: The search strategy was: ("fetus"[MeSH Terms] OR "fetus"[All Fields] OR "fetal"[All Fields]) AND ("electroencephalography"[MeSH Terms] OR "electroencephalography"[All Fields] OR "eeg"[All Fields]) AND (Clinical Trial[ptyp] AND "humans"[MeSH Terms]). Because the landscape of fEEG research has been international, we included studies in English, French, German, and Russian.

Results: From 256 screened studies, 40 studies were ultimately included in the qualitative analysis. We summarize and report features of fEEG which clearly show its potential to act as a direct biomarker of fetal brain health during delivery, ancillary to fetal heart rate monitoring. However, clinical prospective studies are needed to further establish the utility of fEEG monitoring intrapartum. We identified clinical study designs likely to succeed in bringing this intrapartum monitoring modality to the bedside.

Limitations: Despite 80 years of studies in clinical cohorts and animal models, the field of research on intrapartum fEEG is still nascent and shows great promise to augment the currently practiced electronic fetal monitoring.

Prospero Number: CRD42020147474.

Keywords: EEG, labor, fetus, neonates, infant, magnetoencephalogram, electrocorticogram

OPEN ACCESS

Edited by:
Gunnar Naulaers, KU Leuven, Belgium

Reviewed by:
Maria Luisa Tataranno, UMC Utrecht Brain Center, Netherlands
Ana Alarcon, Hospital Sant Joan de Déu Barcelona, Spain

***Correspondence:**
Martin G. Frasch
mfrasch@uw.edu

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *25 June 2020*
Accepted: *07 August 2020*
Published: *11 September 2020*

Citation:
Castel A, Frank YS, Feltner J, Karp FB, Albright CM and Frasch MG (2020) Monitoring Fetal Electroencephalogram Intrapartum: A Systematic Literature Review. Front. Pediatr. 8:584. doi: 10.3389/fped.2020.00584

INTRODUCTION

Perinatally-acquired fetal brain injury is a major cause of long-term neurodevelopmental sequelae, and the single greatest contributor to disability worldwide (1, 2), accounting for 1/10th of all disability-adjusted life years (3). Moreover, intrapartum-related death is the 2nd leading cause of neonatal mortality and the 3rd leading cause of death in children under five (4). Thus, there is an urgent need to identify early signs of fetal distress during labor to allow timely and targeted interventions.

Fetal acidemia contributes to perinatal brain injury (5), and is one of the most common and potentially devastating labor complications. Acidemia occurs in about 25 per 1,000 live births overall and in 73 per 1,000 live preterm births (6, 7) and the risk of subsequent brain injury rises 9-fold in the setting of preterm birth. These risks are even higher with additional complications, such as intraamniotic infection or intrauterine growth restriction (IUGR). Over 90% of children with perinatal brain injury, including that causing cerebral palsy, have a normal life expectancy, but many cannot fully participate in society or fulfill their developmental potential (8).

Today, continuous fetal heart rate (FHR) monitoring is used as an indirect surrogate indicator to suspected fetal acidemia during labor and it fails at that (9). Fetal acidemia *per se* is a poor proxy to fetal brain injury (10). It is then not surprising that FHR monitoring intrapartum does not reliably predict fetal brain injury. Moreover, the fear of missing fetal distress increases the rate of cesarean delivery, with significant maternal risk (11). Meanwhile, about 50% of cesarean sections are deemed unnecessary (12, 13). Conversely, labor is sometimes allowed to proceed when current FHR monitoring technology suggests that the fetus is tolerating it, only to discover later that fetal brain damage occurred, causing a range of signs from subtle neurologic deficits to more overt conditions like cerebral palsy.

Fetal electroencephalogram monitoring intrapartum (fEEG), as a direct monitor of fetal brain activity, was a focus of clinical research as early as the 1940s (14) and into the 1970s (15) and 1990s (16, 17). Notably, Eswaran et al. used a regular FHR scalp electrode and a routine GE HC Corometrics FHR monitoring device to record auditory evoked brainstem potentials, i.e., *evoked* fEEG activity (16). Due to technical limitations and the difficulty of data interpretation, this research into fEEG was not able to be adopted into clinical practice.

The goal of this article is to provide a systematic review of the current literature on intrapartum fEEG. Using the most relevant information gathered from studies on this subject, the second goal of this review is to provide recommendations in order to help ensure successful fEEG study enrollment in the unique environment of a labor and delivery (L&D) unit.

Abbreviations: CRI, continuous rate infusion; CTG, cardiotocogram; fECoG, fetal electrocorticogram; fECG, fetal electrocardiogram; fEEG, fetal electroencephalogram; fMEG, fetal magnetoencephalogram; FHR, fetal heart rate; IUGR, Intrauterine growth restriction; PROM, premature rupture of membrane; UCO, umbilical cord occlusion.

METHODS

The methods for searching and analyzing the relevant literature and for data extraction followed recommendations from the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement. The review has been registered with the PROSPERO database under the number CRD42020147474.

We conducted a literature search in the database PubMed covering all dates and using the following keywords: ("fetus"[MeSH Terms] OR "fetal"[MeSH Terms]) AND ("electroencephalography"[MeSH Terms] OR "electroencephalography"[All Fields] OR "eeg"[All Fields]) AND ("humans"[MeSH Terms]). All the studies retrieved with this search and available in English, French, German and Russian were screened for pertinence by the co-authors who are proficient in these languages. The literature review was completed on April 4, 2020. The eligibility criteria used to determine whether a study was included in this review or not were that the abstract and the full text described the use of EEG on the fetus during labor and provided details about how it was performed. Study selection relied on two reviewers applying the eligibility criteria and selecting studies for inclusion. More specifically, one reviewer screened all the studies and determined if they were relevant or not and the other reviewer examined all the decisions. In case of a disagreement, a joint decision was made upon discussion with the second reviewer. Non-systematic literature reviews were excluded.

The following information was extracted from each study retrieved with the above-mentioned search, and logged in a preformatted spreadsheet: the article name, authors, PubMed identification number, publication year, whether it passed screening or not (1 = passed, 0 = excluded), eligibility (1 = passed or 0 = excluded), the study type (human or animal model), the study size (number of subjects), the gestational age of the subjects when available, the follow-up period if applicable, the electrode configuration, sampling frequency, and monitor type. For excluded studies, the reason for its exclusion was also noted: for those excluded at screening, the reason was categorized and recorded [1 = EEG not mentioned in abstract or article, 2 = irrelevant]. For studies considered non-eligible, the reason was also categorized and recorded [1 = No EEG monitoring, 2 = No information about EEG acquisition or analysis, 3 = EEG done on older children or adults and 4 = fEEG not recorded during labor or fetal magnetoencephalogram (fMEG)].

To present individual study data, quantitative data (such as gestational age) were presented as averages and standard deviations. A Prisma flow diagram was created and all the eligible studies reported in this diagram were reviewed and synthesized.

Each study was classified according to its level of evidence according to the Oxford Center for Evidence-Based medicine level of evidence (18). Level 1 represented a systematic review of inception cohort studies, a systematic review of randomized trials, or n-of-1 trials. Level 2 represented either inception cohort studies, individual cross-sectional

studies with the consistently applied reference standard and blinding, randomized controlled trials or observational study with dramatic effect. Level 3 represented non-consecutive studies, or studies without consistently applied reference standards and non-randomized controlled cohort/follow-up study. Level 4 represented a case series, case-control studies, or poor-quality prognostic cohort study. Finally, level 5 represented expert opinions without an explicit critical appraisal, expert recommendations, or first principles as well as case reports (or case series of less than or equal to 5 cases).

Finally, we summarized systematically individual study findinigs and used information gathered from some of the clinical studies to provide recommendations for successful enrollment in future studies in L&D units.

RESULTS

A Prisma flow diagram showing the results of our database search and presenting the final number of studies included in the systematic review is shown in **Figure 1**.

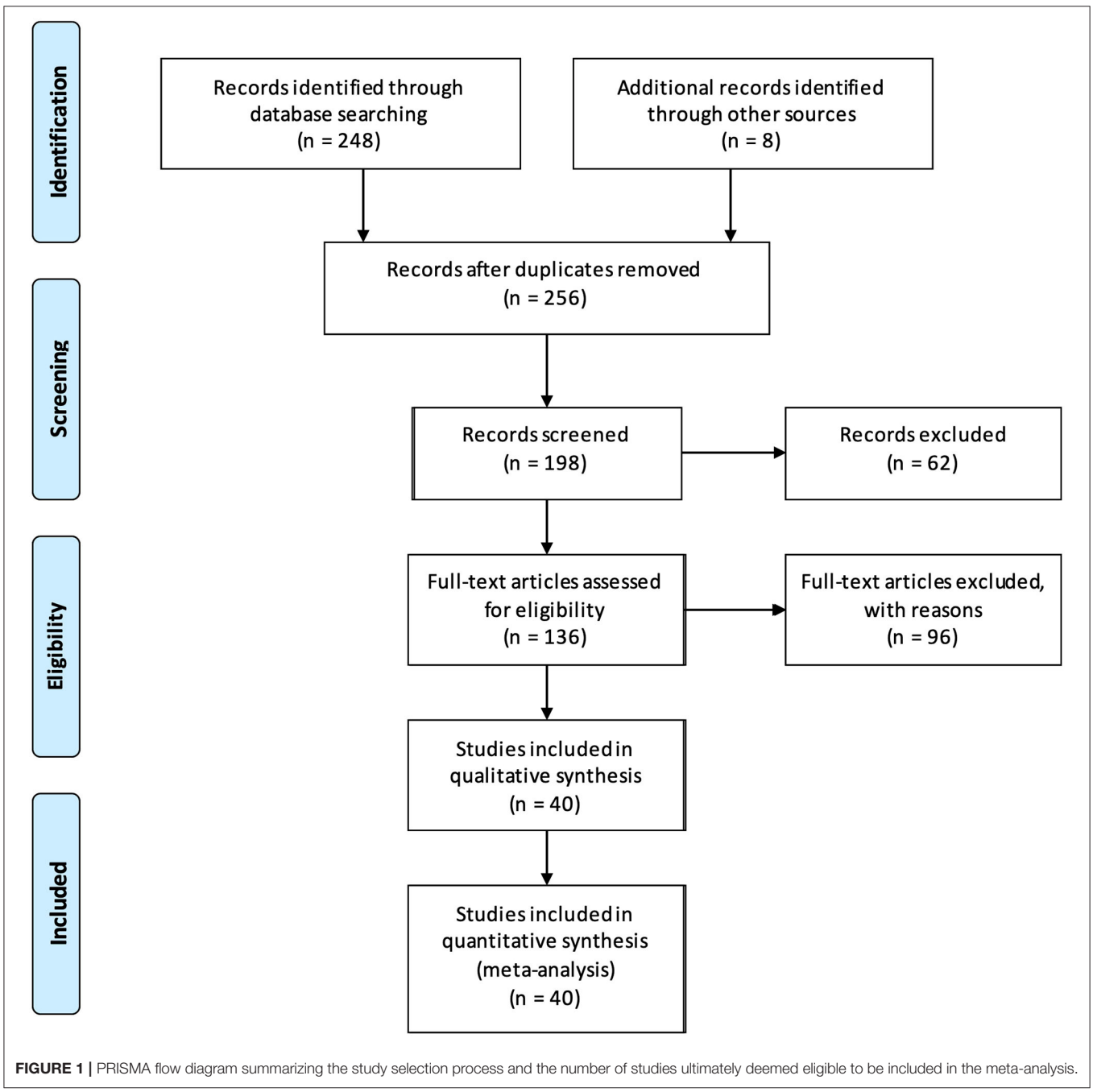

FIGURE 1 | PRISMA flow diagram summarizing the study selection process and the number of studies ultimately deemed eligible to be included in the meta-analysis.

Our initial search yielded 248 results. Eight additional studies were added following cross-referenced review bringing the total number of identified studies to 256. Of these, 34 articles were discarded because they were in a foreign language other than French, German or Russian, 23 additional articles because they were non-systematic reviews and 1 more was excluded because we could not get access to the full text. The initial screening with abstract reviewing was therefore performed on 198 studies and the number of relevant studies was further reduced to 136: 14 additional articles were excluded because there was no EEG performed; the remainder of the studies ($n = 48$) was excluded because they were deemed irrelevant to the subject of our review because the subject of the study did not cover fEEG intrapartum or only mentioned it in passing.

Of these 136 studies, 11 were excluded because they were literature reviews that were not caught in the initial screen and one was an abstract only. The full text was examined for eligibility in the remaining 124 studies.

Of these, 4 studies were excluded because fEEG was not performed as part of the experiment (only as a side test), 16 studies were excluded due to lack of information about the fEEG acquisition or analysis, 22 studies were excluded because EEG (or fMEG) was performed on the fetus but not intrapartum (5 of them were in preterm fetal sheep, 2 in preterm guinea pigs, and the rest in a preterm fetus *in utero*) and 40 studies were excluded because the EEG was performed on neonates after birth, or on older children (i.e., not on fetus or neonates) or adults (the mother). In an additional study, fEEG was studied just before and after labor (but not during), so it was also excluded. One last study was excluded because it was found to be a duplicate from another study written in a different language. Therefore, 40 studies were ultimately included in our analysis.

A summary of the 40 eligible studies is provided in **Table 1**.

Critical Evaluation of the Level of Evidence

Among the 40 eligible studies, none had a level of evidence of 1, 10 studies had a level 2, 11 studies had a level 3, 12 studies had a level 4 and 7 had a level 5. With our search criteria, we identified only a small number of studies with a high level of evidence (i.e., 2 or above), especially studies in humans. In particular, the older studies were mostly either the author's personal experience, case reports, or poor-quality cohort studies as fEEG was in the early experimental stages. However, these studies have the benefits of describing how the technique was developed and perfected over the years to allow determination of normal intrapartum fEEG pattern as well as recognition of patterns that could be indicative of fetal distress. We summarized below the most relevant information gathered from the 40 eligible studies on intrapartum fEEG. We divided them between studies in human fetuses and studies using animal models.

Results of the Individual Studies

Details about the condition under which the fEEG was performed and the monitoring characteristics for all eligible studies are provided in **Table 2**.

Studies in Human Subjects

The first report of fetal EEG was a case report by Lindsley (14) who studied his own child during the 3rd trimester of his wife's pregnancy. For this recording, abdominal probes were used and the tracing had a significant amount of artifacts preventing proper assessment.

Most of the eligible studies in humans date back from the 1960s-70s and more precisely originate from Rosen, Chik, and their team who are among the pioneers of fEEG recording during labor. Several of the findings described in these studies seem to overlap and are summarized below.

Electrodes

The use of fEEG in humans using scalp electrodes during labor was initially reported by Bernstine et al. (59). Later, Rosen and his team perfected the technique (47). A good electrode was defined as: (1) safe to use and easily applied during labor, (2) screening out electrical artifacts such as the movement of the fetal head, maternal movements and the electrical "noise" of uterine contractions, (3) eliminating the electrical pattern of the FHR from the tracing and (4) providing EEG of a technical quality equal to that in the extrauterine environment (49). The group tried different techniques. They initially reported the use of metal skin clips soldered to a shielded cable, coated with non-conductive plastic glue, and filed at their tip to prevent deep scalp penetration (51). This type of electrodes was replaced by cup electrodes, initially with a platinum needle embedded in a lucid disc (49, 51) (with possible skin penetration of 1–2 mm), later replaced by a central silver or platinum pin avoiding penetration of the fetal skin (31). Although this technique seemed to provide reliable and interpretable fEEG signals, artifacts from fetal electrocardiogram (ECG) or movements of the leads remained a common occurrence and these electrodes required continuous suction to stay in place. Mann et al. (45) described the use of a vacuum electrode similar to Rosen et al. but with a silver disc electrode used instead of their platinum needle, thus preventing puncture of the fetal scalp. The main feature of their electrode was the 100% conductivity with a silver cup, wire and plug, low resistance, good suction, and no clogging of the orifices to the vacuum source with the use of a mesh filter (45).

In 1974, Heinrich et al. reported the use of a new intrapartum multimodal fetal monitoring device, the RFT Fetal Monitor BMT-504, that was capable of recording fEEG and tissue oxygen pressure among other parameters (ECG, pressure signals like intraamniotic pressure, temperature, heart rate) combined with either stainless steel clip electrodes or the current standard of care screw electrodes by Corometrics (USA) (35). Weller et al. (19) later described the use of a flexible electrode incorporating a guard ring surrounding the recording sites and forming the indifferent and common electrodes, with the guard ring acting as a short circuit for fECG to prevent its artifact on the EEG tracing. Suction was not needed to maintain in place this type of electrode and its pliability allowed it to be inserted through a 3 cm dilated cervix even if the two electrodes were 23 mm apart. Infrared telemetry was used to display and record fEEG, preventing power line interference, avoiding trailing leads between patient and monitoring equipment, and ensuring

TABLE 1 | Summary of the 40 eligible studies and their level of evidence, number of subjects included and their gestational age.

Authors/year	Level of evidence	Type of study	Number of subjects	Gestational age of subjects
Studies in human				
Thaler et al., 2000 (17)	3	Non-consecutive cohort study	14	39.9 ± 1.2 weeks
Weller et al., 1981 (19)	4	Case series	20	Term fetus
Kurz et al., 1981 (20)	3	Cohort study	20	Not reported
Wilson et al., 1979 (21)	2	Inception cohort study	25	Full term
Chik et al., 1979 (22)	5	Author's recommendations	N/A	N/A
Borgstedt et al., 1978 (23)	4	Poor quality cohort study (biased recruitment of high-risk cases)	158	40.1 ± 2.1 weeks
Nemeadze, 1978 (24)	3	Non-randomized controlled cohort	105	N/A
Chik et al., 1977 (25)	3	Retrospective cohort study	61	39.4 ± 3 weeks
Revol et al., 1977 (26)	4	Case series	140	125 term fetus, 6 near-term and 9 premature between 31 and 36 weeks
Sokol et al., 1977 (27)	2	Prospective cohort study with good follow up	38	Not reported
Chik et al., 1976 (28)	3	Retrospective cohort study	11	Term fetus (mean 40.5 weeks).
Chik et al., 1976 (29)	3	Retrospective cohort study	9	39.1 weeks: 7 term fetus, one 37 weeks and one 34 weeks
Hopp et al., 1976 (30)	4	Retrospective cohort study with poor follow up	85	Not reported
Borgstedt et al., 1975 (31)	2	Prospective cohort study with good follow up	96	Not reported
Chik et al., 1975 (32)	3	Retrospective cohort study	N/A	Not reported
Challamel et al., 1974 (33)	4	Case series	100	92 term fetuses and 8 preterm (<36 weeks).
Fargier et al., 1974 (34)	4	Prospective cohort study with poor follow up	120	Not reported
Heinrich and Seidenschnur, 1974 (35)	5	Proof of concept	1	Not reported
Beier et al., 1973 (36)	2	Prospective cohort study with good follow up	34	Not reported
Carretti et al., 1973 (37)	4	Individual case control study	20	36 to 40 weeks
Hopp et al., 1973 (38)	2	Inception cohort study	37	Not reported
Peltzman et al., 1973 (39)	5	Case series ≤ 5 cases	5	Term fetus
Peltzman et al., 1973 (40)	5	Case series ≤ 5 cases	2	Not reported
Rosen et al., 1973 (41)	4	Poor quality cohort study	6	Not reported
Rosen et al., 1973 (42)	4	Case series	300	Not reported
Chachava et al., 1972 (43)	2	Inception cohort study		Not reported
Hopp et al., 1972 (44)	4	Case series	5	Not reported
Mann et al., 1972 (45)	5	Observation/ first principles	50	Not reported
Feldman et al., 1970 (46)	5	First principles	N/A	N/A
Rosen et al., 1970 (47)	4	Poor quality cohort study	125	Not reported
Chachava et al., 1969 (48)	3	Cohort study	30	Not reported
Rosen and Scibetta, 1969 (49)	5	First principles (technique description)	14	N/A
Barden et al., 1968 (50)	4	Case series	6	Not reported
Rosen and Satran, 1965 (51)	3	Non-consecutive cohort study	15	Not reported
Studies in animals				
De Haan et al., 1997 (52)	2	Individual RCT	21	Fetal lamb: 126.5 ± 2.8 day of gestation (term 147 d)

(Continued)

TABLE 1 | Continued

Authors/year	Level of evidence	Type of Study	Number of subjects	Gestational age of subjects
Thorngren-Jerneck et al., 2001 (53)	3	Exploratory cohort study	16	Near-term fetal lambs at mean (range) gestational age 136 (134–138) days
Kaneko et al., 2003 (54)	3	Exploratory cohort study	8	Fetal lamb: 127–130 days of gestation
Gerrits et al., 2005 (55)	2	Individual RCT	22	Fetal lamb: 117–124 days of gestation
Frasch et al., 2011 (56)	2	Individual cohort study	10	Fetal lamb: 125 ± 1 days gestation
Wang et al., 2014 (57)	2	Individual cohort study	20	Near term fetal lamb: 123 ± 2 days

electrical safety. Artifacts due to the movement of leads were also prevented by incorporating the first stage of amplification in the composite assembly thus avoiding long wires carrying low-level signals. With this device, artifact-free fEEG recordings were obtained 80% of the time and uterine contraction did not affect the signal. However, in the case of a breech presentation, no fEEG could be recorded.

Problems and Limitations of the Technique

The two major problems associated with intrapartum fEEG precluding its routine use were technical issues and data interpretation. Placement of electrodes over the occipital area is the area where electrodes are most easily applied but because the occiput is a relatively quiet electrical area of the brain, the parietal area is preferred (51). Because of the limited space, there is only a limited number of electrodes that can be placed precluding comparison of homologous areas of the brain (51). Additionally, the moist scalp and uterine environment can attenuate potentials. Therefore, isolating the scalp from the environment by using suction allowed the recording of higher amplitude potentials.

Failure to obtain adequate EEG tracing was reported to most often occur when the signal was obscured by fECG (49). Simultaneous recording of fEEG and fECG was shown to aid in the recognition of ECG artifacts (**Figure 2**) (44). Thankfully, newer electrodes were later developed to help limit the number of artifacts from fECG and movements (19). Finally, the use of infrared telemetry to transfer the fEEG to display and recording equipment helped to prevent power line interference as previously mentioned (19).

Another initial limitation of the technique was the amount of information that needed to be visually interpreted. Indeed, visual interpretation had significant methodologic and interpretation bias and required certain expertise preventing routine use of fEEG as part of intrapartum monitoring. Evaluation of the value of digitized minute-to-minute and even second-to-second fluctuation in fEEG amplitude and frequency was reported by Peltzmann et al. (39). These authors used a computer system to extrapolate the mean baseline fEEG line crosses (per 5-s epochs) as well as the mean integrated fEEG amplitude and presented the data in graphs with plotted point corresponding to the calculated mean by 5-s epochs. This represented the first steps toward simplification and standardization of fEEG signal analysis.

At the same time, computer algorithms were created to facilitate fEEG signal interpretation and standardize their evaluation (32). A computer program was developed to help fEEG analysis by replacing the cumbersome visual analysis in an effort to integrate fEEG in computer-assisted intrapartum data management and was shown to provide 85–95% consistency with visual interpretation (32). This program classified fEEG patterns as Low Voltage Irregular, Mixed, High Voltage Slow, Trace Alternant, Voltage Depression, Isoelectricity, and Artifact.

Visual processing of the fEEG in the form of the spectral display as an adjunct to digital analytic technique to reduce the ambiguity in fEEG interpretation was initially described by Peltzmann et al. (40). Years later, Kurz et al. (20) described the use of spectral power analysis performed in 30 s intervals with the results plotted continually over the course of the entire observation in waterfall style. The authors suggested that continuous fEEG spectral power plotting helped detect artifacts on the fly which still occur and must be dealt with during the interpretation of the fEEG patterns.

Thaler et al. (17) reported the use of real-time spectral analysis to monitor fEEG during labor as more objective analysis of fEEG signal. Real-time Fast Fourier Transform algorithm allowed the representation of the EEG signals in terms of the relative power of the various frequencies of which it is composed. These frequencies were then displayed by using a density spectral array technique which helps visualize the contribution of each frequency band to the overall power spectrum: delta (0.3 to 3 Hz), theta (4 to 7 Hz), alpha (8 to 11 Hz), sigma (12 to 14 Hz) and beta (15 to 32 Hz). The brightness of a given pixel represented the relative power present at the corresponding frequency element in the fEEG. A spectral time record appeared as a black and white or grayscale image in which a given spectrum would take up only a single row of pixels. In addition, the display of the Spectral Edge Frequency (SEF) indicated the highest dominant frequency of the fEEG signal (i.e., the frequency below which 90% of the spectral power resides).

More recently, our team created automated algorithms for unsupervised fEEG-FHR monitoring and for the detection of fEEG-FHR patterns pathognomonic of adaptive brain shut-down

TABLE 2 | Summary of the method used for fetal electroencephalogram (or electrocorticogram) recording and the condition of recording for the 40 eligible studies.

Authors/year (ref)	Electrode type and placement	Mode/frequency	Condition of recording
Studies in human			
Thaler et al., 2000 (17)	Two custom-made circular scalp EEG electrodes (suction silicone rubber cups applied by continuous negative pressure) with a central metal probe applied at the occipitoparietal or parietal region (with at least 4 cm between electrodes). FHR recorded with a scalp electrode.	Signals sampled at 250 Hz, stored and displayed by a Cerebro-trac 2500 (*SRD Medical Ltd. Shorashim, Israel*) using real-time Fourier transform (FFT) algorithm to calculate the power spectrum of fEEG. Epochs length acquisition: 4 s. Band-pass filter: 1.5–30 Hz. Amplifier sensitivity: 200 μV.	Low risk pregnancies in the active stage of labor.
(20, 58)	*N/A*	FHR, uterine contractions, fetal blood analysis and fEEG recorded polygraphically. Spectral power analysis was performed in real time sequentially and plotted in 30 s intervals continually over the course of the entire observation in waterfall style.	Normal deliveries ($n = 20$)
Weller et al., 1981 (19)	Flexible electrode minimizing ECG artifact with an incorporated guard ring surrounding the recording sites and forming the indifferent and common electrodes (acting as a short circuit to the fECG). The two electrodes are 23 mm apart and inserted through a 3 cm dilated cervix (after membrane rupture).	Amplifier circuit: microminiature resistors and standard low noise operational amplifier (SF C 2776UC). Input resistance: 2.5 + 2.5 Ohms. Gain: 32 dB. Power requirement ± 5 volts (100 μA). Noise level set at < 2 μV peak to peak (1 to 40 Hz). Infrared telemetry used to convey the fEEG to display and record equipment.	Monitoring during the second phase of uncomplicated labor in primigravid mothers with term fetus under epidural anesthetic. Recording in a standard delivery room.
Wilson et al., 1979 (21)	8 channel portable Elena Schonander Recorder.	Continuous recording. Fetal EEG analyzed in 10 s epochs.	High-risk African primigravida mother.
Chik et al., 1979 (22)	*N/A*	*N/A*	*N/A*
Borgstedt et al., 1978 (23)	Same as Rosen et al., 1973b (42)	Same as Rosen and Scibetta, 1970	Selected high risk cases (based on prenatal maternal complication or suspected intrapartum fetal distress)
Nemeadze, 1978 (24)	Simultaneous fetal EEG and ECG recording from fetal head from the moment of the first stage of labor when cervical dilation was 4-6 cm. No further details provided.	N/A	Study of the impact of premature rupture of the membrane in a group of healthy women ($n = 60$) and a group of women with mild nephropathy ($n = 45$). fEEG and fECG are recorded simultaneously during labor and correlated to onset of PROM and the Apgar score.
Chik et al., 1977 (25)	2 scalp cup electrodes held by applied suction on fetal head with a central silver or platinum pin avoiding penetration of the fetal skin.	Same as Rosen and Scibetta, 1970	File selection of high-risk infants monitored for at least 1 h during labor.
Revol et al., 1977 (26)	2 scalp cup electrodes held by continuous suction with a central silver pin, placed on the parietal region of the skull.	Continuous recording during labor and after. EEG device: Mingograf EEG 8 Siemens (ink jet print). Filter from frequencies >30 Hz. Band-pass with 0.15 s time constant. Calibration 1 s, 50 μV.	125 cases: normal labor (35), abnormal labor (18), fetal distress (19), under anesthesia (alfatesine $n = 21$) or ketamine ($n = 26$).
Sokol et al., 1977 (27)	Same as Rosen et al., 1973b PMID: 4681833	Same as Rosen and Scibetta, 1970	EEG recorded in suspected increased risk deliveries. EEG findings assessed in relation to follow up at 1 year.
Chik et al., 1976a (28)	Same as Rosen et al., 1973b	Same as Rosen and Scibetta, 1970	Same as Rosen et al., 1973b
Chik et al., 1976b (29)	Same as Rosen et al., 1973b	Same as Rosen and Scibetta, 1970	Same as Rosen et al., 1973b
Hopp et al., 1976 (30)	Same as Hopp et al., 1972 (44)	Same as Hopp et al., 1972 (35)	Simultaneous recording of fEEG and CTG in 220 fetuses, (85 cases ultimately included). with some under conditions of intermittent hypoxia due to uterine contractions and after maternal administration of drugs.

(Continued)

TABLE 2 | Continued

Authors/year (ref)	Electrode type and placement	Mode/Frequency	Condition of recording
Borgstedt et al., 1975 (31)	2 scalp cup electrodes held by applied suction on the fetal head with a central silver or platinum pin avoiding penetration of the fetal skin. Electrodes implanted once cervix dilation reached 2 cm.	Same as Rosen and Scibetta, 1970	Patients selected because of increased prenatal risk or suspected fetal distress during labor.
Chik et al., 1975 (32)	Same as Rosen et al., 1973b	Same as Rosen and Scibetta, 1970	Retrospective evaluation of fEEG recording.
Challamel et al., 1974 (33)	3 Modified Dassault electrode (cup with central pin) placed on the scalp with one placed in parietal position and a minimal distance between electrodes of 4 cm.	EEG device: Mingograf EEG 8 Siemens (ink jet print). Filter from frequencies >30 Hz. Band-pass with 0.15 s time constant. Calibration 1 s, 50 μV. Continuous recording (30 min to 5 h intrapartum). Best derivation exploited.	Monitoring during different conditions of labor (including fetal distress).
Fargier et al., 1974 (34)	Same as Challamel et al., 1974	Same as Challamel et al., 1974. No filter <70 Hz.	Monitoring during different conditions of labor and different drug administrations.
Heinrich and Seidenschnur, 1974 (35)	Same as Rosen et al., 1973 (42)	Intrapartum multimodal fetal monitoring device: RFT Fetal Monitor BMT-504.	One example of EEG recorded in a neonate to show how the new monitor can be used.
Beier et al., 1973 (36)	1 cup electrode consisting of a central needle pin surrounded by a 4 cm disc and two suction grooves connected with a suction device to ensure firm electrode placement over the fetal skull.	1-channel EEG; simultaneous recording of fetal ECG, FHR and intraamniotic pressure channels. Calibration 1 s, 50 μV.	About 30 min duration intrapartum monitoring in 34 fetuses from healthy pregnancies, labor and postnatal outcomes. In 15 fetuses/neonates, the corresponding postnatal recordings were also made.
Carretti et al., 1973 (37)	Plexiglas suction cup with 6 electrodes around its periphery placed on the fetus occiput following >4 cm cervix dilation.	Galileo apparatus allowing multipolar EEG recording.	Comparison of fEEG in healthy mother before and after oxygen administration.
Peltzman et al., 1973 (39)	2 flexible stainless-steel screw electrodes (impedance <400 Ohms in all cases)	FEEG is continuously monitored on a polygraph (Grass instrument Co., Quincy, Massachusetts). Band-pass filter: 1.0 to 35.0 Hz. A PDP-7 computer allows real time analysis and storage of the fEEG before the information is transmitted to a 64-channel analog/digital converter. The fEEG for each 5 s epochs is set to zero mean. Mean fEEG and time integrated fEEG amplitude are also computed. The program then computes a zero line-cross count on the zero-mean fEEG by checking along the wave within each 5 s epoch and recording the line cross each time the polarity changes. Data based upon 12 to 19 continuous 5 s epochs are presented as graphs with plotted points from the beginning of the recording and represent artifact-free analog fEEG.	Uneventful labor and delivery (3 out of 5 were induced labor). Effect of analgesia with meperidine ($n = 3$) and diazepam ($n = 2$) or paracervical block ($n = 2$) evaluated.
Peltzman et al., 1973 (40)	2 flexible stainless-steel bipolar screw electrodes placed 3–4 cm apart of the parietal area of the fetal scalp.	Similar to Peltzman et al., 1973	Uneventful labor with paracervical block (1% mepivacaine) administered
Rosen et al., 1973a (15)	Same as electrodes as Rosen et al., 1970 Electrodes placed along the sagittal suture and between the fontanelles to avoid the forceps blade in forceps birth (over the parietal region for spontaneous birth)	Similar to Rosen and Scibetta, 1970	FEEG during expulsion efforts and during low forceps deliveries are compared to EEG during spontaneous deliveries. Pre-forceps EEG also recorded for comparison.
Rosen et al., 1973b (42)	2 electrodes applied over the parietal and consisting of a silver pin in the center of a lucite disc maintained in place by continuous suction after application with an interelectrode distance of at least 4 cm.	Similar to Rosen and Scibetta, 1970	EEG recording during different labor situations (eventful vs. complicated labor in neurologically abnormal infants).

(Continued)

TABLE 2 | Continued

Authors/year (ref)	Electrode type and placement	Mode/Frequency	Condition of recording
Hopp et al., 1973 (38)	Same as Hopp et al., 1972 (44)	Same as Hopp et al., 1972 (44)	Simultaneous recording of fetal EEG, fECG and CTG in 37 fetuses during labor in the late first stage and in the stage of active pushing (second stage).
Chachava et al., 1972 (43)	Two fEEG and fECG electrodes placed onto the fetal head, as remotely from each other as possible, fixated using vacuum suction.	N/A	fEEG, fECG and maternal EEG recorded following cervical dilation and rupture of membranes during the first stage of labor.
Hopp et al., 1972 (44)	3 cup electrodes made of a 40 mm silver disc with a 5 mm center pin are used: two biparietal and one on midline, placed and held by suction.	Fetal ECG, CTG and intra-amniotic pressure acquired with an 8-channels EEG device. Artifact free fEEG was achieved when bimodal rejection mode was chosen (i.e., biparietal EEG electrode against a ground).	Simultaneous recording of fEEG and CTG during different labor conditions including normal labor and cardiac decelerations.
Mann et al., 1972 (45)	Two silver disc electrodes consisting of a vacuum contact cup, tube and wire connection and a vacuum electric plug to the vacuum module are placed with membranes ruptured after cervix dilation reaches 3–4 cm. The electrodes are placed up against the vertex about 1 to 2 cm apart.	Serial bipolar EEG are obtained by direct write out on the dual-channel San'Ei electroencephalograph (San'Ei Instrument Co., Div. of Medical System Corp, Great Neck, New York)	Description of the electrode used to record fetal EEG during labor (50 recordings). Specific conditions of recording not documented.
Rosen and Scibetta, 1970 (47)	Two electrodes consisting of a platinum needle embedded in the center of a lucite disc with the firm margin of the disc preventing deep penetration of the needle. Continuous suction is applied within the disc to draw the scalp up to the needle recording point (with possible skin penetration of 1–2 mm). Circular grooves into the lucite disc prevent the skin from occluding the suction. Electrodes implanted as soon as cervical dilation reached 3 cm over the parietal areas.	Bipolar EEG recording using an 8-channel Dynograph Recorder run at 30 mm/s. Time constants are arranged to allow recording wave frequencies between 0.5 and 32 Hz/min. Recording amplitude of 20 μV/cm and 50 μV/cm are used.	125 fetuses recorded during different conditions (normal labor, forceps assisted delivery and following drug administration).
Chachava et al., 1969 (48)	Two fEEG scalp electrodes were held with vacuum suction and placed 2–3 cm apart on the fetal head (after 3–4 cm cervical dilation and ruptured membranes).	Bipolar EEG recorded with either a 4-channel locally made device or an 8-channel EEG device by Orion.	fEEG recorded intrapartum: 20 with normal labor and 10 with complications.
Rosen and Scibetta, 1969 (49)	Two electrodes made of an outer shell in silicone rubber with its periphery circumscribed by a silicon rubber guard ring impregnated with powdered silver (the outer ring is a patient ground) and a platinum needle sheathed in a Teflon tube soldered to a silver plate wire with its distal 2 mm left bare, are used. The electrodes are introduced after membranes ruptured once cervix dilation reached 3 cm. Suction is turned on after the electrode is applied on the fetal head.	Recording in standard EEG fashion with filters admitting wave frequencies between 0.5 and 32 Hz. Paper speed is 30 mm/s recording amplitude at 20 to 50 μV/cm. To document the EEG activity as brain waves, the technique of evoked response to a 34 dB, 35 ms sound (repeated every 4 s) is used.	Fetal EEG during labor (no more detail provided).
Barden et al., 1968 (50)	Skin-clip electrode placed on the presenting vertex.	A summing computer (Mnemotron Corporation CAT-Model 400 B). is used to accentuate fEEG response time locked to an acoustic signal and to cancel non time-locked fECG and random electrical noise signals. FEEG responses were sequentially averaged.	Elective induced labor. FEEG recorded before, during and after the onset of a 1,000 Hz, pure tone of 450 s duration (88 to 105 dB).
Rosen and Satran, 1965 (51)	Metal skin clips soldered to shielded cable, coated with non-conductive plastic glue and filed at their tip to prevent deep scalp penetration and attached to the vertex with a modified uterine packing forceps. Mother grounded to the machine by a strap around the thigh.	Grass Model III portable EEG (Grass Instrument Co., Quincy, Mass).	Normal labor condition studied.

(Continued)

TABLE 2 | Continued

Authors/year (ref)	Electrode type and placement	Mode/Frequency	Condition of recording
Studies in animals			
De Haan et al., 1997 (52)	Two pairs of EEG electrodes (AS633-5SSF, Cooner Wire Co., Chatsworth, CA) placed on the parasagittal fetal dura through burr holes (skull coordinates relative to bregma: anterior 5 mm and 15 mm, lateral 10 mm).	The total fEEG intensity is median filtered to remove short-term (<20 min) fluctuations, log transformed to get a better approximation of the normal distribution and normalized with respect to the 12-h baseline. Total EEG intensity, EEG spectral edge (upper 90% of frequency), and cortical impedance are measured in 15-min periods during UCO and in 5-h intervals after the last occlusion. Epileptiform activity and spike detection software (Monitor, Stellate Systems, Quebec, Canada) is used to scan the raw EEG file.	FEEG recorded after UCO with fetuses randomized to one of three groups: group I, repeated total UCO for 1 min every 2.5 min; group II, repeated total UCO for 2 min every 5 min; and group III, no occlusions (sham controls). UCO is repeated until fetal arterial blood pressure had fallen below 2.7 kPa (20 mm Hg) during two successive occlusions, or until fetal blood pressure failed to recover to baseline levels when the next occlusion is due.
Thorngren-Jerneck et al., 2001 (53)	Two EEG electrodes (shielded stainless steel) placed bilaterally over the parietal cortex (10 mm anterior of bregma and 15 mm lateral of midline), inserted through drilled holes in the parietal bone. A subcutaneous reference electrode is placed posteriorly in the midline of the skull.	No detail provided.	16 near-term fetal lambs: 8 lamb fetuses exteriorized and subjected to total UCO in a water bath, four lamb fetuses exteriorized and serving as sham controls and four lamb fetuses immediately delivered after minimal preparation and serving as healthy controls
Kaneko et al., 2003 (54)	Electrodes of Teflon-coated stainless-steel wire (Cooner Wire, Chatsworth, Calif) implanted biparietally on the dura for recording of electrocortical activity.	No details provided.	ECoG recording following repeated UCO for 4 min, every 90 min, and over 6 h (total 4 UCO).
Gerrits et al., 2005 (55)	Two pairs of EEG electrodes (AS633-5SSF; Cooner Wire Co., Chatsworth, CA) placed on the dura over the parasagittal parietal cortex (5 and 15 mm anterior and 10 mm lateral to the bregma), with a reference electrode sewn over the occiput.	Fetal parietal EEG and impedance recorded continuously. Signals are averaged at 1-min intervals and stored to disk by custom software (Labview for Windows; National Instruments Ltd, Austin, TX), running on an IBM compatible computer. The EEG signal is low pass filtered at 30 Hz, and the intensity spectrum and impedance signal are extracted. The raw EEG signal is recorded for off-line detection of seizure events.	Fetal lamb subjected to selective cooling of the head following cerebral ischemia (with one control group)
Frasch et al., 2011 (56)	Stainless steel ECoG electrodes are implanted biparietally on the dura through small burr holes in the skull bone placed ∼1–1.5 cm lateral to the junction of the sagittal and lambdoid sutures. The bared portion of the wire to each electrode is rolled into a small ball and inserted into each burr hole to rest on the dura with a small plastic disk covering each burr hole held with tissue adhesive against the skull bone. A reference electrode is placed in the loose connective tissue in the midline overlying the occipital bone at the back of the skull.	ECG and ECoG are recorded and digitized at 1,000 Hz. For ECG, a 60 Hz notch filter is applied. For ECoG, a band pass 0.3–30 Hz filter is used. The ECOG signal is sampled down to 100 Hz prior to analysis. Voltage amplitude and 95% spectral edge frequency (SEF), are calculated over 4 s intervals (Spektralparameter, GJB Datentechnik GmbH, Langewiesen, Germany).	Fetal lamb studied after series of mild, moderate and severe UCO until fetal arterial pH fell below 7.00
Wang et al., 2014 (57)	Fetal instrumentation after exteriorization: a modified FHR electrode with a double spiral placed on the fetal head is used.	A PowerLab system is used for data acquisition and analysis (Chart 5 For Windows, AD Instruments Pty Ltd, Castle Hill, Australia). For fEEG recording, a band pass 0.3–30 Hz filter were used. Prior to analysis, fECoG and fEEG were sampled down to 100 Hz.	FEEG (and fECoG) recorded in near term fetal lamb during repeated UCO.

fECG, fetal electrocardiogram; fECoG, fetal electrocorticogramm; fEEG, fetal electroencephalogram; UCO, umbilical cord occlusion.

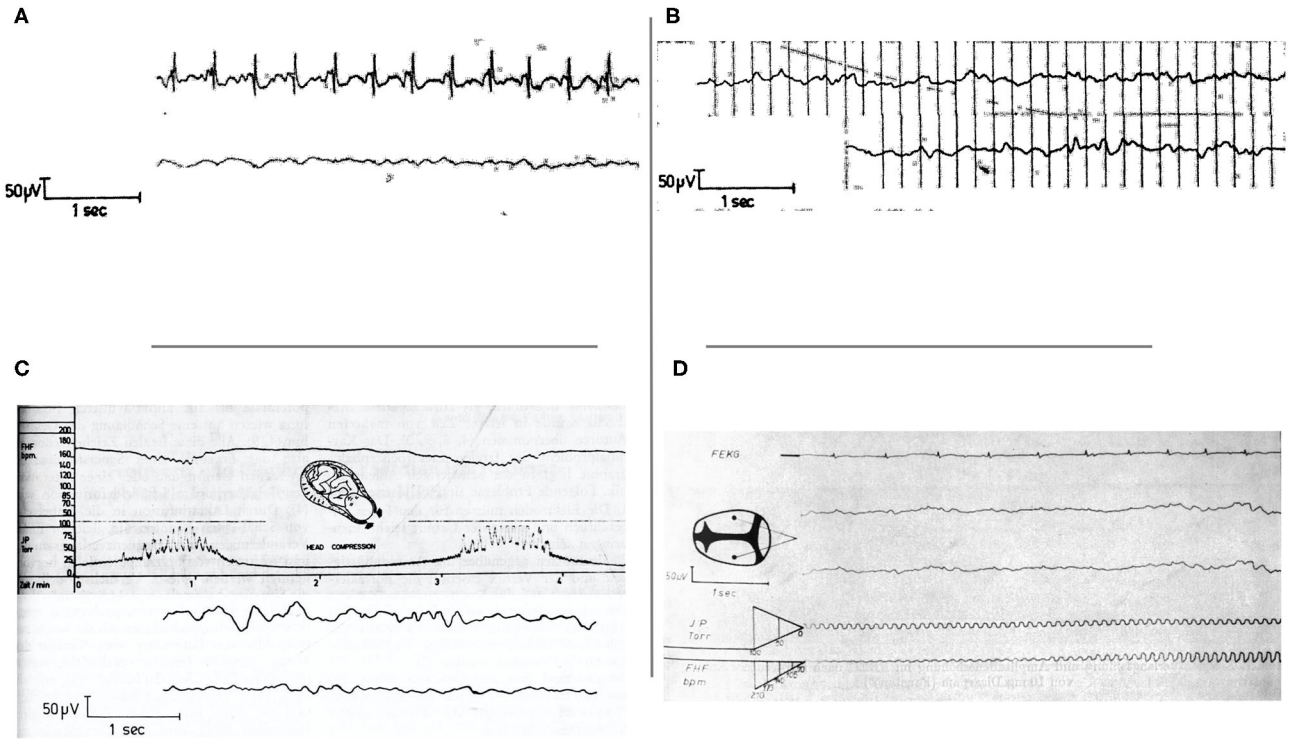

FIGURE 2 | (A) Simultaneous recording of fECG and fEEG. Artifacts from fECG effect on fEEG can be identified by recording both traces simultaneously. From Hopp et al. (44). **(B)** Intra and post-partum fetal/neonatal EEG recordings showing the great similarity between both traces. From Hopp et al. (44). EEG, electroencephalogram. **(C)** Cardiotocogram (top) and fEEG (bottom) recorded during early cardiac deceleration. The fEEG pattern represents the change during contractions with high amplitude low-frequency waves and the recovery once the contractions ceased. From Hopp et al. (44). fEEG, fetal electroencephalogram. **(D)** Simultaneous recording of fECG (top trace), two-channel fEEG (middle two traces), and FHR (bottom trace). This figure shows fEEG changes during severe variable deceleration. The fEEG trace shows waves of low amplitude and near isoelectricity as well as intermittent spike potentials between contractions. From Hopp et al. (30). fECG, fetal electrocardiogram; fEEG, fetal electroencephalogram; FHR, fetal heart rate.

as an early response to incipient acidemia and cardiovascular decompensation (57).

Intrapartum EEG Findings

The early studies described the fEEG signal observed during labor under different conditions and while most of them were initially just observations, they allowed to gain the experience needed to determine what a normal fEEG during labor should look like and what should be interpreted as abnormal (15, 26, 33, 36, 42, 45, 49, 51, 60).

To document the fEEG activity, the technique of evoked response can be used (49), although results can be quite unpredictable with significant artifacts (50).

A summary of fEEG findings associated with normal labor, abnormal labor, and following drug administration is presented in **Table 3**.

1. *EEG findings during normal labor*

During labor, a low voltage baseline pattern is noted (51). The study of 14 acceptable fEEG revealed that the voltages varied from 5 to 50 μV/cm and the wave frequencies were found between 1 and 25 Hz (49). A small change in electrical activity was noted after delivery and rarely low voltage (20 μV), faster (8 *per second*) waves compared to the fEEG trace seen 30 s after delivery, and not seen before were observed after the umbilical cord was clamped. On most tracings, the electrical activity before and after the first breath and before and after the cord was clamped did not appear to change abruptly (34, 51). As the recording continued, the electrical activity slowly increased in voltage and approached that seen in similar brain regions in neonates. About 5 min after delivery, the tracing could not be distinguished from the tracing of alert neonates several hours old. Rosen and Satran (51) concluded that fEEG activity recorded early in labor has a baseline pattern similar to that of the alert neonate. Studies of 125 additional fEEG by the same team confirmed that the fEEG patterns observed during normal labor were similar to those present in neonates of the same weight. The wave frequencies varied between 0.5 and 25 Hz with the predominant frequencies in the 2.5–5 Hz (47). Similarly, Chachava et al. (48) reported fEEG findings during 20 normal labors and found that healthy (physiological) fEEG was characterized by low-amplitude waves of 0.04–2 s duration which the authors note was within the range of the reported spectrum of antenatal fEEG frequencies observed (0.5–30 Hz according to Humar & Jawinen as well as according to

TABLE 3 | Summary of the possible findings in fetal EEG tracings, as reported in the different studies, under different labor conditions.

Fetal EEG conditions	Observations
Baseline fEEG during normal labor	Rosen and Satran, 1965: Low voltage baseline pattern with a low voltage (20 μV), faster frequency (8 Hz) after umbilical cord clamping. On most tracings, the electrical activity before and after the first breath and before and after umbilical cord clamping did not appear to change abruptly. FEEG activity recorded early in labor has a baseline pattern similar to that of the alert neonate (51). Chachava et al., 1969: Low amplitude waves (10–30 μV) of 0.04–2 s in duration with observation of alpha, beta, theta and delta waves (48). Rosen and Scibetta, 1969: The fEEG voltages vary from 5 to 50 μV per cm with waves frequencies between 1 and 25 Hz. Rosen et al., 1970: The wave frequencies vary between 0.5 and 25 Hz with the predominant frequencies in the 2.5–5 Hz. Patterns similar to those present in neonates of the same birth weight (60). Hopp et al., 1972: Amplitude of fEEG ranges between 10 and 70 μV. Frequency varies considerably and ranges between 2 and 20 Hz (44). Mann et al., 1972: Rhythm consisting of 1 to 3 Hz waves with an amplitude of about 40 to 75 μV and superimposed faster frequencies of 4 to 8 Hz and 10 to 30 μV (45). Fargier et al., 1974: Similar to that of neonates of same gestational age with presumptive alternance of awake/sleep states (both deep and active sleep) (34). Borgstedt et al., 1975: Wave frequencies of 0.5 to 25 Hz with an amplitude generally between 50 and 100 μV/cm similar to neonatal EEG (31). Thaler et al., 2000: Two fundamental EEG patterns are identified: high voltage slow activity (HVSA) (quiet behavioral state) and low voltage fast activity (LVSA) (active behavioral state). On average, LVSA was present 60.1% of the time and HVSA was present 39.9% of the time (17).
Contractions	Mann et al., 1972: No fEEG changes (even with very intense oxytocin-induced contractions) (45). Rosen et al., 1973: No fEEG changes even with stronger contractions during the second stage of labor (15, 42). Beier et al., 1973: No fEEG patterns associated with contractions in most cases, but in some fetuses, a 20–40 s delayed increase in amplitude which normalized 5–10 s after contractions ended is observed (36). Hopp et al., 1976: fEEG shows reduction of frequency and increase in wave amplitude during contractions (30, 61). Chachava et al., 1972, Fargier et al., 1974, Revol et al., 1977, Weller et al., 1981: No fEEG changes with normal contractions (19, 26, 34, 43).
Spontaneous birth	Rosen et al., 1973b: Low voltage irregular activity. Artifactual distortion of the fEEG baseline characterized by large rolling waves of almost 2 s in duration due to electrodes movements when the vertex moves rapidly and the fEEG is recorded in the microvolt range (42).
Abnormal situations	
Fetal heart rate (FHR) decelerations	Rosen et al., 1970: Previously recorded higher voltage fEEG pattern abruptly changed in association with depression of the FHR to an almost flat or baseline pattern. The tracing returned to pre-existing patterns after FHR returned to normal. This change was found to be most commonly associated with delayed FHR decelerations (60). Hopp et al., 1972: Fetal bradycardia, especially during contraction-associated late decelerations, was accompanied by reduction in fEEG waves (lower frequency) and occurrence of fEEG spike potentials (44). Rosen et al., 1973b: No change with early decelerations. With variable and late deceleration, the fEEG appeared to lose the faster rhythms, followed by a more apparent slowing. Then, isoelectric to almost flat periods with rare bursts of fEEG are seen and finally a totally isoelectric interval might be observed sometimes for longer than 10 s (rarely more than 30 s). As the FHR returns to its baseline rate, the reverse of this progression takes place. The entire sequence from onset to return may last from 30 s to longer than 1 min (42). Fargier et al., 1974: No changes with early deceleration (34). Revol et al., 1977: Early deceleration was only associated with fEEG changes only with FHR below 90 bpm (26). Hopp et al., 1976: During severe variable decelerations, fEEG showed waves of low amplitude and near isoelectricity and intermittent spike potentials between contractions (30, 61). Wilson et al., 1979: A significant relationship was noted between the increasing percentage of electrocerebral silence and the development of FHR deceleration patterns during labor. Early FHR deceleration was also associated with prolonged silence in the fEEG (21).
Tachycardia	Rosen et al., 1970: fEEG changes consistent with voltage suppression, i.e., generalized decrease in the wave amplitude of a constant nature often associated with increasing intervals of flattening without EEG activity (60). Thaler et al., 2000: FHR accelerations typically associated with periods of low voltage slow activity (17).
Fetal distress	Chachava et al., 1969: fEEG of a baby born asphyxiated and demised within 15 min postpartum showed fast activity around 6 Hz that may represent brain hypoxia. High amplitude low-frequency waves were suspected to be signs of brain injury during labor (48). Hopp et al., 1973: Patterns pathognomonic for abnormal fEEG and suspected fetal brain injury: (1) extremely high voltage activity (>80 μV), (2) extremely low voltage activity (< 10 μV), (3) spike potentials as a sign of epileptiform activity, (4) bihemispheric differences, (5) Reduction of fEEG frequency during pathologically silent FHR pattern (38). Rosen et al., 1973b: Non-transient fEEG changes such as sharp waves were defined as repetitive waves, always of the same polarity, generally higher in amplitude than the surrounding fEEG and generally <50 ms in duration. When observed, they were usually present at the onset of recording and continued throughout labor. These sharp waves seemed to be more frequent in the developmentally abnormal child at 1 year of life. The combination of sharp waves and low voltage did not occur in the normal population. Therefore, this type of activity may suggest fetal distress (42). Borgstedt et al., 1975: Isolated sharp waves were more frequent in newborn with abnormal neurologic findings than in those neurologically normal (31).

(Continued)

TABLE 3 | Continued

Fetal EEG conditions	Observations
	Revol et al., 1977: Cases with the combination of abnormal fEEG, abnormal cerebral blood flow, low pH (<7.25) and abnormal Apgar score had the lowest Apgar score. FEEG did not always normalize after *in utero* resuscitation despite the correction of FHR and pH (26). Borgstedt et al., 1978: Prolonged voltage suppression (<20 μV) (during the entire recording or for at least several seconds, alternating with normal trace and continuing through birth) is correlated with lower 1 min- and 5 min- Apgar score and the need for post-partum resuscitation (23). Wilson et al., 1979: Significant correlation between the development of electrocerebral silence in the fEEG and the development of fetal acidosis. The rapid deterioration in fEEG occurred as the pH fell and even at preacidotic levels (pH of 7.2 to 7.25) marked changes were present with the cessation of electrical activity in the fetal brain (21). Nemeadze, 1978: No effect of PROM on fEEG, but pregnancy history compounds fEEG response to PROM as follows. The brief fEEG period in response to a PROM event including the immediate 5-min post-PROM represents an adaptive functional response of the fetal brain indicative of fetal reserve. Under normal delivery, this response subsides within 1–3 min post-PROM and is associated with a healthy birth (in the absence of any other delivery complications). In contrast, prolongation of this recovery period post-PROM to 4–5 min indicates a reduction in fetal adaptation capability and was associated with brain injury at birth (24).
Head compression from cephalopelvic disproportion.	Wilson et al., 1979: Head compression did not appear to influence fetal brain activity (21).
Uterine hypertonia or hyperkinesis	Challamel et al., 1974, Fargier et al., 1974, Revol et al., 1977: Decreased activity and flattening of the fEEG signal (26, 33, 34)
Forceps	Rosen et al., 1973a: Forceps application was not associated with any change, but traction was, with an almost flat fEEG tracing observed (15, 62). Challamel et al., 1974, Fargier et al., 1974: High forceps extraction compared to low forceps extraction was always associated with fEEG changes during the traction phase and characterized by flattening of the trace returning to normal after a few seconds if the extraction was short and not too intense. Repeated and prolonged tractions were associated with persistent isoelectric trace up to the birth of the child with changes persisting for 20 min after (33). Revol et al., 1977: During forceps or vacuum extraction, fEEG changes that disappear just before the expulsion effort or persistent fEEG changes sometimes to the point of isoelectric trace were noted in all but one case (26).
Oxygen administration to the mother	Carretti et al., 1973: FEEG changes within 1.5–2 min after initiation of O_2 are characterized by a progressive increase in amplitude and frequency of the waves (from 1–5 Hz to 8–12 Hz) reaching a maximum at 7–8 min followed by a decrease in the activity of the trace to return to baseline activity after 12–15 min in half of the cases (37).
Drugs	
Meperidine	Rosen et al., 1970: Early responses: a transient increase in delta wave frequencies (2.5–5 Hz), about 50 μV in amplitude first seen between 1 and 2 min after IV injection of the drug followed by trace alternance-like pattern of bursty activity with 5 min after the mother was given the medication. This pattern could last as long as 2 h after the injection. As the time interval after injection increased, the presence of faster, lower voltage forms (5–10 μV) and (15–25 Hz) in the beta range became more obvious (47, 60). Peltzman et al., 1973a: No identified fEEG changes (39, 40).
Ketamine	Fargier et al., 1974: Development of sharp theta activity on an initially normal baseline, then progression to fewer waves and flattening of the trace to the point of isoelectricity with occasional bursts of theta activity (34).
Pethidine	Hopp et al., 1976: One min post injection of a 50 mg dose, there is a reduction of amplitude and frequency of fEEG activity. These changes are more pronounced 4 min post injection. At 6 min post injection, resynchronization is observed. These effects persisted for 25 min post injection and fEEG normalized more or less within 105 min post injection (30, 61).
Barbiturate	Fargier et al., 1974: Sodium thiopental: same as ketamine as well as small high-frequency low voltage waves on a normal baseline progressing to decreased activity and flattening of the trace (34). Revol et al., 1977: Significant changes with long periods of isoelectric traces (26).
Local anesthesia	Rosen et al., 1970: With local carbocaine, transient increase in higher voltage (50 μV/cm) bursty waves (15–25 Hz) was noted. These changes appeared to be transient (47, 60). Peltzman et al., 1973a and Peltzman et al., 1973b: Decrease in fEEG amplitude (39, 40). Challamel et al., 1974: Local epidural with marcaine or bupivacaine associated with a high-frequency rhythm (when using a filter frequency >30 Hz) with clusters of rhythmic theta waves (33).
Penthrane (anesthetic gas)	Rosen et al., 1970: Trace alternant picture persists while the gas is being administered during the terminal stages of labor (47, 60).
Diazepam	Peltzman et al., 1973a: No identified fEEG changes (39, 40). Hopp et al., 1976: The fEEG frequency decreased within 30 min post injection of 10 mg of diazepam and the amplitude increased to 80 μV. The EEG was normal when recorded in the neonate 40 min after the injection (30, 61).
Less mature infant with analgesic medications	Rosen et al., 1973b: Persistence of all voltages below 20 μV with prolonged intervals of isoelectricity (low voltage tracing) associated with an initially normal amplitude and pattern of recording that then changes to persistent low voltage with prolonged periods of isoelectricity (42).
Long-term outcome	
Infants with normal long-term outcome	Chik et al., 1976a: Study of fEEG of children neurologically normal at 1 year of age: the mixed pattern was predominant accounting for 41.2% of the epochs. Trace alternants accounted for 32.2%, high voltage slow for 21.5% and low voltage irregular for 4.4% of the patterns. Less than 0.2% showed depression or isoelectricity. In the neonatal EEG studied, there was a decrease in the relative frequency of mixed and an increase in high voltage slow patterns (28, 29).

(Continued)

TABLE 3 | Continued

Fetal EEG conditions	Observations
Infants with abnormal long-term outcome	Borgstedt et al., 1975: Sharp waves that appear in isolation and not as part of burst activity were identified as abnormal (31). Chik et al., 1976b: Study of fEEG of children neurologically abnormal at 1 year of age: the relative frequency of the low voltage irregular pattern is increased with a consequently decrease in mixed and high voltage slow patterns. The mean relative frequency of the low voltage irregular pattern was significantly greater with the lower Apgar scores (<9). Therefore, low voltage irregular activity occurred more frequently in the neurologically abnormal group (compared to the neurologically normal group) (28, 29). Sokol et al., 1977: The detection of the presence of sharp waves alone on fEEG allowed to correctly classify 76% of patients in terms of the neurological outcome at 1 year (normal vs. abnormal). Detection of voltage depression alone appropriately identified 68% patients. The observation of a combination of sharp waves and/or prolonged voltage depression improved the identification of abnormal infants but misclassified half of the normal infants as abnormal (66% correct). These results confirm the relationship between sharp waves and voltage depression in the fEEG and abnormal infant outcome (27).

bpm, beat per minute; EEG, electroencephalogram; fEEG, fetal electroencephalogram; FHR, Fetal heart rate; Hz, Hertz; s, seconds.

Bernstine and Borkowski) (59, 63) They reported an amplitude of 10–30 μV with the observation of alpha, beta, theta, and delta waves.

Hopp et al. (44) reported simultaneous acquisition of fEEG and cardiotocogram (CTG) during labor using 3 scalp electrodes (2 biparietal and 1 midline). Normal fEEG was characterized by an amplitude ranging between 10 and 70 μV with high variability in frequencies ranging between 2 and 20 Hz. They also concluded that fEEG and neonatal EEG are basically identical and could not be differentiated from each other (**Figure 2**). Borgstedt et al. (31) also reported normal fEEG showing wave frequencies of 0.5 to 25 Hz with an amplitude generally between 50 and 100 μV/cm, similar to neonatal EEG (31). Three studies from the same group reported similar findings with intrapartum fEEG showing alternance of active and quiet sleep phase similar to neonates (26, 33, 34).

In a study by Mann et al. (45), adequate fEEG were obtained and studied in 50 patients. The EEG prior to, during and following a very intense contraction (∼95 mmHg after oxytocin infusion) was characterized by a rhythm consisting of 1 to 3 Hz waves with an amplitude of about 40 to 75 μV with superimposed faster frequencies of 4 to 8 Hz and 10 to 30 μV. There were no significant changes in the fEEG signal during the uterine contraction and this fEEG was very similar to that of the same patient examined 18 h after birth. The lack of influence of uterine contractions or expulsion on the fEEG signal was also documented by Chachava et al. (43) and Challamel et al. (33). Similarly, fEEG recorded during the second stage of labor did not show any alteration in frequency, amplitude, and pattern despite the increase in uterine pressure associated with maternal pushing (a contraction of abdominal wall muscles) (15). Conversely, in a study using simultaneous CTG and fEEG recording under conditions of intermittent hypoxia due to uterine contractions, fEEG showed a reduction of frequency and increase of wave amplitude during contractions (30).

During spontaneous birth, a low voltage irregular activity was noted as well as artifactual distortion of the fEEG baseline characterized by large rolling waves of almost 2 s in duration due to electrodes movements when the vertex moves rapidly and the fEEG is recorded in the microvolt range (15). This appears to be a common problem during the birth process.

The effect of head compression associated with cephalopelvic disproportion on fetal brain activity was studied and no significant differences in fEEG findings between the group with cephalopelvic disproportion and the group without it were noted (21).

Using real-time spectral analysis, a more objective method of fEEG assessment, Thaler et al. (17) identified two fundamental fEEG patterns in the recording: high voltage slow activity (HVSA) (quiet behavioral state) and low voltage fast activity (LVFA) (active behavioral state). FHR accelerations were typically associated with periods of LVFA but there was no relationship between uterine contractions and SEF or density spectral array (DSA) (power spectrum). The 90% SEF was found to be an excellent index of cyclic EEG activity. When combining the results of the 14 fetuses, it was found that on average, LVFA was present 60.1% of the time and HVSA was present 39.9% of the time.

2. *Abnormal EEG findings*

Chachava et al. (48) first reported fEEG during complicated labor and presented the case of a baby born asphyxiated and demised within 15 min postpartum. The fEEG showed fast activity around 6 Hz that was suggested to represent brain hypoxia but the changes were not considered unique and pathognomonic. High amplitude low-frequency waves were, in their experience, signs of intrapartum brain injury.

During labor, transient or persistent fEEG changes can be observed. Usually, persistent changes are considered to be abnormal if they occur between two events such as uterine contraction or expulsion efforts leading to a progressive deterioration of the fEEG activity (26).

Evaluation of the fEEG signal associated with FHR changes revealed different situations which we summarized below.

Simultaneous recording of fEEG and CTG/FHR revealed that slow waves and frequency decrease could be observed during and shortly after uterine contractions and were seen as an expression of short-term brain ischemia due to an increase in intracranial pressure (**Figure 2**). The vagal stimulation inducing the early decelerations in CTG was thought to be due to an increase

of intracranial pressure, but indirectly, with the primary vagus stimulation trigger being due to transient cerebral hypoperfusion during uterine contraction. Fetal bradycardia, especially during contraction-associated late decelerations, was accompanied by a reduction in fEEG waves (lower frequency) and occurrence of fEEG spike potentials (44).

In one study by Rosen et al. (42), transient fEEG changes were noted during FHR deceleration. The fEEG appeared to lose faster rhythms, followed by a more apparent slowing. As the condition persisted, isoelectric to almost flat periods with rare bursts of fEEG were seen. Finally, a totally isoelectric interval was observed sometimes for longer than 10 s (rarely more than 30 s). As the FHR returned to its baseline rate, the reverse of this progression took place with the entire sequence from onset to return lasting from 30 s to sometimes longer than 1 min. These changes were not seen with early FHR deceleration but were observed with variable decelerations and late decelerations. They were also observed during prolonged spontaneous expulsion or expulsion of a distressed infant (33). In another study, using simultaneous CTG and fEEG recording, severe variable decelerations were also associated with waves of low amplitude and near isoelectricity and intermittent spike potentials between contractions (**Figure 2**) (30).

Revol et al. (26) documented the fEEG changes during spontaneous expulsion and noted the following events: either no fEEG changes or transient fEEG changes not exceeding 20 s (Type I); EEG changes that disappear just before the expulsion effort (Type II); and persistent fEEG changes (Type III). During expulsion, the relationship between fEEG and FHR revealed that Type I fEEG was 85% of the time associated with transient or no FHR changes whereas Type II and Type III fEEG changes were associated with bradycardia. In their cases, early deceleration was only associated with fEEG changes with FHR below 90 bpm (26). Spontaneous tachycardia (>160 bpm) and bradycardia were both associated with fEEG changes (decreased activity and flattening of the trace) (34). However, tachycardia following atropine administration was not associated with any fEEG changes (34). Using spectral power analysis, (20) also observed a similar relationship between the degree of spectral fEEG suppression and the FHR decelerations induced by uterine contractions (20).

Simultaneous recording of fEEG and fECG during labor after premature rupture of membranes (PROM) in a group of healthy women and a group of women with mild nephropathy, was reported by Nemeadze (24). Normal fEEG characteristics were 1–16 Hz, 10–30 μV, asynchronous, dysrhythmic activity; PROM had no significant effect on these parameters of fEEG or fECG over the course of labor. However, the immediate response of fEEG to the PROM event, identified as a period of 5 min, was shown to be reflective of incipient perinatal brain injury. Specifically, in some fetuses, regardless of maternal health status, PROM induced a response in fEEG characterized by pathological activity with high-amplitude slow 1–3 Hz waves, periodically acquiring a group-rhythmical character; fEEG normalized gradually within 1–3 min and fetuses showing this rapid recovery of fEEG all had a healthy birth. Fetuses whose EEG recovered within 3–5 min and belonging to mothers with mild nephropathy, however, were diagnosed with brain injury at birth. It is thus concluded that pregnancy complications, but not the PROM itself, impact the acute fEEG response to PROM and may provide valuable insights into therapeutic labor management.

The effect of forceps birth on fetal brain activity was also evaluated (15, 26, 33, 34, 51). FEEG recorded during labor involving forceps application required placement of the two electrodes along the sagittal suture and between the fontanelles to avoid the forceps blade (compared to their placement over the parietal region for normal birth) (15). Aperiodic, 50 μV, 0.5–5 Hz slow waves were reported to become more apparent when forceps was applied or when the vertex was on the perineum and the mother bore down (51). Forceps application was not associated with any changes in the fEEG signal but during traction, an almost flat tracing was observed. Tracing resembling a burst suppression pattern could also be observed in some cases (15). Another study comparing high and low forceps extraction revealed that high forceps extraction was always associated with fEEG changes during the traction phase and was characterized by flattening of the trace returning to normal after a few seconds if the extraction was short and not too intense (33). Repeated and prolonged tractions were associated with persistent isoelectric trace up to the birth of the child and sometimes persisting for at least 20 min after birth (33, 34).

A significant correlation between the development of electrocerebral silence in the fEEG during the final hour of the first stage of labor and the development of fetal acidosis at the end of the first stage of labor was reported (21). The rapid deterioration in the fEEG occurred as the pH fell and even at preacidotic levels (pH of 7.2 to 7.25), marked changes were present with the cessation of electrical activity in the fetal brain. A significant relationship was also noted between the increasing percentage of electrocerebral silence and the development of FHR deceleration patterns during labor. In the study by Wilson et al. (21), different from Rosen et al. (15), early FHR deceleration was associated with prolonged silence in the fEEG. Similarly, intermittent suppression of fetal brain electrical activity during FHR decelerations induced by umbilical cord occlusions and also arising at around pH values of 7.2 was reported in fetal sheep models of human labor (56, 64).

Another fEEG study in 11 cases of fetal distress revealed a loss of fEEG variability, sometimes similar to the awake state. These changes were usually transient during events or maneuvers (33). A decrease in fEEG amplitude and frequency has been reported with the uterine hypertonicity of hyperkinesia (26, 33). Revol and his team studied fEEG changes associated with fetal distress in 37 cases (fetal distress diagnosed with a combination of abnormal scalp pH, umbilical blood pH, and Apgar score at 1 min). In 4 additional cases, some changes in fEEG were suspicious for fetal distress. The fEEG was abnormal in 39 of these 41 cases. The 26 cases for which all the aforementioned criteria of fetal distress were present had the lowest 1 min Apgar score (between 1 and 7). In 8 of these cases, *in utero* resuscitation measures allowed improvement of biological (i.e., pH) values and FHR. However, only in 2 cases did the fEEG normalize before birth which supports a delay in fEEG recovery compared to other criteria (26). Another study looked at the correlation between abnormal fEEG findings and the 1 min and 5 min Apgar scores

in high-risk cases (23). Prolonged voltage suppression periods (below 20 μV), usually present from the beginning of the fEEG recording and persisting throughout, represented a distinctive pattern significantly correlated with a low 1 min and 5 min Apgar scores. This pattern was also correlated with the employment of postpartum resuscitative measures and with the degree of resuscitation (23).

In a study by Hopp et al. (38), simultaneous evaluation of fEEG, fetal ECG, and CTG, during the first and second stages of labor was shown to improve the detection of the fetus at risk of brain injury. They reported a series of patterns pathognomonic for abnormal fEEG: (1) extremely high voltage activity (> 80 μV), (2) extremely low voltage activity (< 10 μV), (3) spike potentials as a sign of epileptiform activity, (4) bihemispheric differences, and (5) reduction of fEEG frequency during a pathologically silent FHR pattern (44).

During the same period, Rosen and his team also reported one major fEEG abnormality, the non-transient sharp waves defined as repetitive waves always of the same polarity, generally higher in amplitude than the surrounding fEEG and generally <50 ms in duration (42). When observed, they were usually present at the onset of recording and continued throughout labor and seemed to be more frequent in children neurologically abnormal at 1 year of age. This observation was later confirmed by retrospective fEEG evaluations to see if the infant outcome at 1 year of age with regard to neurological status could be predicted (27, 31). Sharp waves that appeared in isolation and not as part of burst activity were identified as abnormal. Once again, isolated sharp waves were noted to be more frequent in newborns with abnormal neurologic findings than in those neurologically normal and were significantly associated with neurological abnormalities at 1 year of age (27, 31).

Furthermore, retrospective comparison of intrapartum fEEG from neurologically abnormal infants at 1 year of age to neurologically normal children revealed that the combination of sharp waves and low voltage did not occur in the normal population suggesting that this type of activity may indicate fetal distress requiring intervention (27, 42) To further confirm these findings, the previously described computer program developed by Chik et al. (32) was used to retrospectively evaluate artifact-free EEG of these neurologically normal and neurologically abnormal infants (25, 28, 29). In the neurologically normal group, the mixed pattern was predominant accounting for 41.2% of the 10,511 epochs evaluated. The trace alternant pattern accounted for 32.2%, high voltage slow pattern for 21.5%, and low voltage irregular pattern for 4.4% of the patterns. Less than 0.2% showed depression or isoelectric signal. In the neurologically abnormal infants, low voltage irregular activity accounted for 17.85% of the epochs, mixed activity for 30.5%, high voltage slow activity for 18.1%, and trace alternant for 33.2%. Less than 0.2% of the epochs showed depression or isoelectric signals. The number of observed fEEG patterns in abnormal cases was significantly different from normal cases. The relative frequency of the low voltage irregular pattern was increased with a decrease in mixed and high voltage slow patterns. The mean relative frequency of low voltage irregular pattern was significantly greater in the 1 min lower Apgar score (<9). Therefore, low voltage irregular patterns were shown to occur more frequently in the neurologically abnormal group (compared to the neurologically normal group). The same group used a computer-interpreted EEG to try to predict the infant neurological outcome at 1 year. Using fEEG patterns alone (by looking at the relative frequency of low voltage irregular, high voltage slow, mixed, and trace alternant patterns), almost two-thirds of the neurologically normal infants and of the abnormal infants were correctly classified. Using intrapartum fEEG and FHR patterns simultaneously provided slightly better results to predict neurologically normal infants but gave the same results for the neurologically abnormal ones. Combining intrapartum data with postpartum data, including 1 min, 5 min Apgar scores, and neonatal neurologic examinations, about 80% of the infants were correctly classified (25, 65). These results show that combining multiple methods of peripartum fetal monitoring allows better detection of fetal distress that could affect long term neurological outcome.

3. *Effect of drugs*

Six studies reported their observations of fEEG following maternal general anesthesia with different drugs and described some characteristic changes (26, 30, 33, 34, 39, 47).

FEEG recorded following maternal anesthesia with alfatesine at a continuous rate infusion (CRI) showed changes between 1 to 11 min (mean 3.5 min) following the beginning of the CRI. Initially, theta waves occurring in clusters altering the baseline rhythm were noted. These fEEG changes eventually disappeared to the point of reaching a discontinuous aspect with alternance of theta wave clusters and isoelectric state. Electrical silence could also be observed. Theta activity was noted to persist for about 30 min after birth. The fEEG baseline activity reappeared about 40 min after birth with the persistence of occasional theta activity during different vigilance states associated with anesthesia. FEEG changes were more pronounced if fetal distress was also present.

The effect of ketamine on fEEG showed similar changes with sharp theta activity on an initially normal baseline with a progression to fewer waves and flattening of the trace to the point of isoelectricity with occasional bursts of theta activity approximately within 3 min following drug administration (26, 34). Barbiturates such as sodium thiopental were associated with the more significant changes with long periods of isoelectric traces (33, 34). Meperidine and diazepam were not found to be associated with any fEEG changes in a very small case series (39). Conversely, meperidine was associated with early fEEG changes characterized by a transient increase in delta and theta wave frequencies (2.5–5 Hz), about 50 μV in amplitude, first seen between 1 and 2 min after intravenous injection of the drug followed by a trace-alternant-like pattern of bursty activity within 5 min after the mother was given the medication (47). This pattern could last as long as 2 h after the injection. These results suggested a rapid transfer of the drug from the mother to the fetus. As the time interval after injection increased, the presence of faster, lower voltage forms (5–10 μV, 15–25 Hz) in the beta range would become more obvious (47). The effect of the administration of 50 mg of pethidine revealed a reduction of

amplitude and frequency of fEEG activity about 1 min after the injection (30). These changes were more pronounced at 4 min post-injection. At 6 min post-injection, resynchronization was observed. These effects persisted for 25 min and fEEG normalized more or less within 105 min post-injection.

Minor fEEG changes were noted with local anesthesia and were characterized by high-frequency rhythms with clusters of rhythmic theta waves (33). FEEG changes associated with anesthetic persisted for 1 to 3 days after birth. In two very small case studies, paracervical block with 1% mepivacaine was associated with a decrease in fEEG amplitude with a questionable effect on the frequency (39, 40). Caudal or paracervical carbocaine administration was shown to produce transient pattern changes consisting of an increase in higher voltage (50 μV/cm) bursty waves (15–25 Hz) (47). In the presence of penthrane, a trace alternant picture persisted while the gas was being administered during the terminal stages of labor (47). Finally, the effect of diazepam injection (10 mg) on fEEG was also reported. The fEEG frequency decreased within 30 min post-injection and the amplitude increased to 80 μV with normalization of neonatal EEG recorded 40 min after the injection (30).

Of interest, the persistence of all recording voltages below 20 μV with prolonged intervals of isoelectricity, described as low voltage tracing, was observed in less mature infants in the presence of analgesic medications (42). This pattern was associated with an initially normal amplitude and the pattern of recording changing to persistent low voltage with prolonged periods of isoelectricity.

Finally, in a study comparing fEEG before and after oxygen (O_2) administration by mask to 20 mothers during labor, it was shown that O_2 administration caused fEEG changes within 1 min 30 s to 2 min after initiation of O_2 characterized by a progressive increase in amplitude and frequency of the waves (from 1–5 Hz to 8–12 Hz) reaching a maximum at 7–8 min followed by a decrease in the activity of the trace to return to baseline activity after 12–15 min in half of the cases (37).

Studies Using Animal Models

All the animal studies deemed eligible used a fetal sheep model. Because of the similarities between ovine and human fetal physiology (66), this species is considered a reliable model to study fetal cerebral development (67). First, the sheep fetus displays cerebral hemodynamics similar to that in humans. Second, the sheep fetal cardiovascular and EEG data can be derived in the unanesthetized state. Third, similar to the human fetus (68–71), the sheep fetus displays a very limited range of cerebral autoregulation under normal conditions and they both have a pressure-passive cerebral circulation when subjected to systemic hypoxia and the associated hypotension (72–76). Such hypotensive response is amplified in chronically hypoxic pregnancies, such as with IUGR, where fetal myocardial glycogen reserves are more rapidly depleted under conditions of umbilical cord occlusions (UCO) (77, 78).

The 6 studies selected used transient UCO mimicking what can happen during labor with uterine contractions and therefore represent a good model compromise to study intrapartum fetal distress and fEEG. Some of these studies recorded fetal electrocorticogram (fECoG) where electrodes are placed directly on the dura for optimal signal quality by comparison with fEEG where electrodes are sewed into fetal sheep's skin.

De Haan et al. (52) recorded fECoG, allowing to record brain electrical activity similar to fEEG in sheep fetuses following repeated UCO of different duration (1 min every 2.5 min or 2 min every 5 min) compared to sham controls. During the occlusions, there was a progressive fall in fEcoG intensity, more pronounced in the group with the longer UCO. FEcoG activity at the final occlusion and recovery to normal sleep cycling patterns were similar in the two UCO groups. A fall in SEF during UCO followed by rapid normalization during recovery was similar in the two asphyxiated groups. Two characteristic patterns of electrophysiologic changes were noted. In the baseline period, there was normal sleep cycling characterized by an alternation of high voltage and low voltage fEcoG activity. During the occlusions, the fECoG intensity decreased to eventually reach a trough at the final occlusion and recovered thereafter. In fetuses that subsequently developed only selective neuronal loss as assessed on histologic evaluation, the fECoG rapidly recovered, associated with very little epileptiform or spike activity. Conversely, fECoG tracing indicating more epileptiform activity was seen in the fetuses with most extensive neurological damage on histology and the fECoG recovery was slower in the more severely damaged fetuses. In comparison, sham fetuses showed no changes in fECoG activity (and at postmortem evaluation). Despite a similar frequency of the asphyxia periods, the longer episodes of cord occlusion appeared to have a greater initial effect on the fEcoG with significantly more epileptiform and spike activity than the shorter one reflecting the cumulative effect of intermittent ischemia with longer hypotensive periods on fECoG (and consequently fEEG) and brain injury.

Thorngren-Jerneck et al. (53) also reported the effect of UCO on fEEG and compared the fEEG signal of 3 groups: one subjected to total UCO until cardiac arrest, one sham control group and one healthy control group. The fEEG became rapidly flat during the cord occlusion in all lambs subjected to UCO and remained isoelectric during the 4 h after delivery. Conversely, the fEEG was "normal," i.e., showing continuous activity with mixed frequencies, in sham and healthy controls during the 4 h after delivery. Using positron emission tomography, they also demonstrated that global cerebral metabolic rate was significantly reduced 4 h after fetal asphyxia induced by UCO. Their findings suggest that prolonged isoelectricity identified on EEG after birth is an indication of severe fetal distress and that a reduction in the brain's metabolic rate represents an early indicator of global hypoxic cerebral ischemia.

In another study by Kaneko et al. (54), fetal sheep were subjected to UCO without regard to the electrocortical state activity every 90 min, and over 6 h (for a total of four UCOs). The fECoG was monitored continuously and assessed by visual analysis into periods of high voltage (>100 μV) and low voltage (<50 μV). Following UCO, an indeterminate electrocortical pattern became apparent with initially lower than baseline electrocortical state and then gradually increasing toward a high voltage electrocortical state but with no evident cycling. The

fetal electrocortical activity was disrupted markedly by 4 min of UCO, with an abrupt flattening of the fECoG. With the release of the cord occluder, the fECoG amplitude increased steadily over several minutes. These results show that UCO resulted in a progressive decrease in fECoG amplitude with most animals showing a flat ECoG by 90 s but with rapid recovery in voltage amplitude after the release of the occluder. These results are similar to what has been reported in humans following severe cardiac deceleration (21, 42).

Our team conducted several studies using a fetal sheep model of human labor and showed that certain changes in fEEG and fECoG accurately predicted severe acidemia during labor with sufficient lead-time to potentially intervene and perform a cesarean section (56, 57, 64, 79). This method was shown to have a positive predictive value (PPV) of 70% and a negative predictive value (NPV) of 100% (57). Using ECoG and EEG recordings, we identified pathognomonic changes in fetal electrocortical activity predictive of cardiovascular decompensation and severe acidemia allowing early (∼60 min) recognition of a critical situation and giving sufficient time to perform an emergency cesarean section (56, 57).

The utility of joint fEEG-FHR monitoring is based on the consistent emergence of synchronized UCO-triggered blood pressure, and fEEG-FHR changes, prior to reaching a severe level of fetal acidemia where brain injury might occur. The fetal blood pressure showed a pathological hypotensive

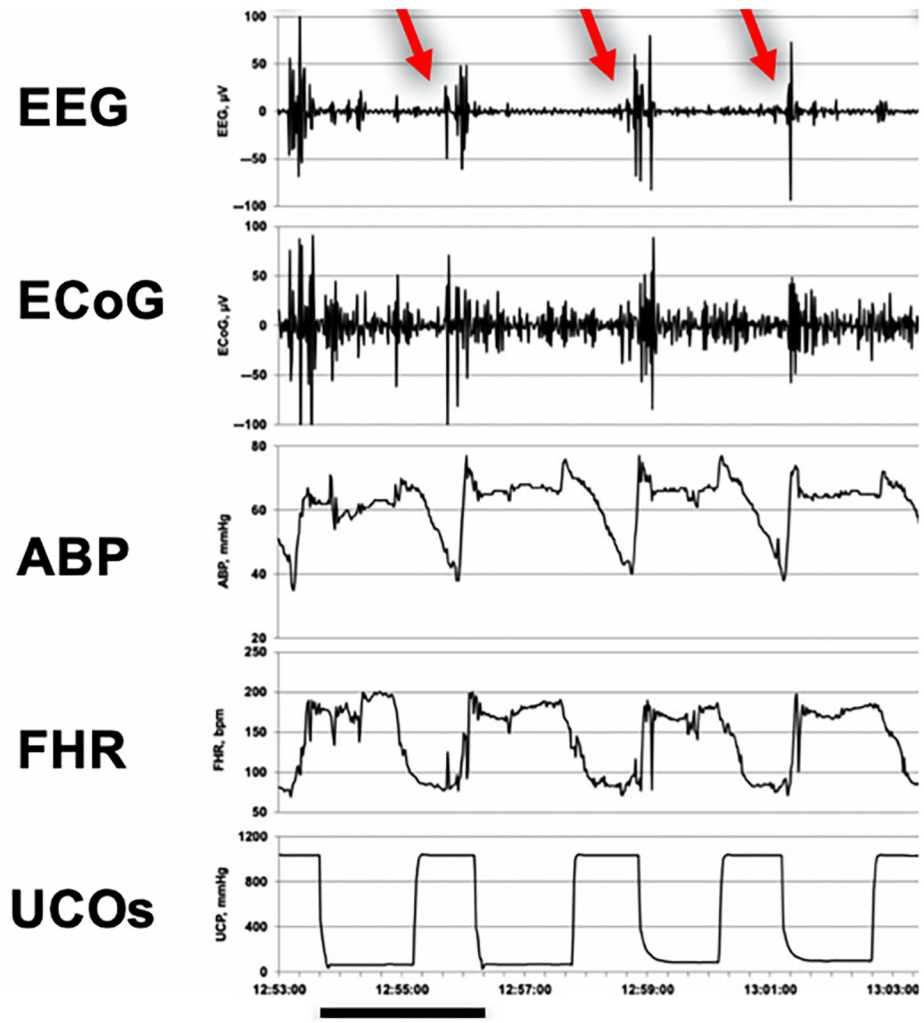

FIGURE 3 | Emergence of EEG-FHR pattern in a fetal sheep model. A representative 10 min recording made during the early stage of severe umbilical cord occlusions (UCOs) at a pH of about 7.2 and about 60 min prior to pH dropping to <7.00 indicated cardiovascular decompensation (hypotensive fetal systemic arterial blood pressure; ABP) in response to FHR deceleration triggered by UCO. It shows the pathognomonic fEEG pattern (black bar = 2.5 min). Red arrows indicate the pathognomonic burst-like EEG activity correlated in time to the FHR decelerations and pathological ABP decreases during the UCOs. UCOs continued until pH < 7.00 was reached in each fetus (about 4 h). Fetal arterial blood samples were taken each 20 min. This timing corresponds to pH of 7.20 seen in 20% of births (62). From Wang et al. (57) EEG, electroencephalogram, μV; ECoG, electrocorticogram, μV; ABP, fetal systemic arterial blood pressure, mmHg; FHR, fetal heart rate, bpm; UCOs, umbilical cord occlusions, mmHg (rise in occlusion pressure corresponds to an UCO).

behavior concomitant with the fEEG-FHR changes during FHR decelerations (**Figure 3**) (57). These changes are thought to be due to adaptive brain shut-down, triggered at a pH of about 7.20. Of note, Yumoto et al. also reported a pH of 7.20 to be the critical value, below which fetal myocardial contractility begins to decrease (80). Adaptive brain shutdown prevents the brain from passing from upper to lower ischemic flow thresholds (81, 82). When the fetal brain blood flow falls beneath the lower ischemic flow threshold, permanent neurological injury occurs (56).

Finally, the chronically instrumented non-anesthetized fetal sheep model with UCO was also used to study the presence of epileptiform activity during rewarming from moderate hypothermia, one of the undesirable outcomes associated with this common therapy for HIE (55). Cerebral ischemia was induced by transient carotid occlusion corroborated by the onset of an isoelectric fEEG signal within 30 s of occlusion. Sheep fetuses were randomized to either cooling or sham cooling starting at 6 h after ischemia and continued until 72 h. Rebound electrical seizure events were observed in about 50% of the cooled animal and 7% of the sham-cooled animals. These results demonstrated that following a severe ischemic insult treated with moderate cerebral hypothermia, rapid rewarming was associated with a significant but transient increase in EEG-defined seizure events.

Taken together, these findings, similar to what has been reported in humans, further emphasize the relevance of the fetal sheep model to study labor-associated fetal and neonatal cerebral ischemia and help develop and validate new monitoring and therapeutic interventions.

Synthesis of Results

The systematic analysis of the literature on intrapartum fEEG remains relatively scarce and somewhat outdated with a lot of redundant or confirmatory information. However, studies in human patients, corroborated by studies using animal models suggest that this monitoring modality can provide valuable information about fetal brain activity that significantly influences and predicts the neurological development of the newborn (28, 29, 31, 57).

One of the key features of fEEG is the ability to potentially detect cerebral activity changes secondary to fetal distress sooner than with evaluation of FHR alone and more continuously than by relying only on fetal scalp blood pH, a technique hardly used in the modern practice (21, 55, 56). If the technical difficulties associated with electrode placements have been mostly removed (19), the problem of objective data analysis and interpretation, although improved by the use of computer algorithms (28, 29, 32, 57) and spectral analysis (17, 20, 39, 40) remains a significant limiting factor in democratizing the use of intrapartum fEEG as part of the routine labor monitoring. Despite compelling evidence that joint fEEG and FHR monitoring and detection of pathognomonic patterns associated with fetal distress are key features of intrapartum fetal health assessment, the development of methods allowing unsupervised monitoring of these two variables without requiring a high level of expertise, remains in its infancy.

Furthermore, as most human studies were either retrospective cohort studies or case series, more clinical prospective studies are needed to further establish the utility of fEEG monitoring intrapartum. We identified clinical study designs likely to succeed in bringing this monitoring modality as a bedside test in the unique setting of L&D and will be discussing them below.

Risk of Bias Across Studies

To limit the risk of bias for each individual study, we ought to assess the studies at the outcome level. However, because the majority of the eligible studies, particularly the ones in humans, reported mainly descriptive findings, this turned out to be extremely challenging. Indeed, a lot of these studies just described fEEG traces of selected cases (19, 22, 26, 33, 39, 40, 42, 49–51, 60). In fact, only 4 studies analyzed the fEEG in relation to the outcome at 1 year and are from the same group (with the same cohort for all but one study) (27–29, 31).

We did try to limit bias in study selection by not just including studies in English, but also those in French, German and Russian which added 30 studies to the screening process with 14 ultimately deemed eligible.

DISCUSSION

Summary of Evidence

The review of the aforementioned eligible studies allowed us to establish some key-points about intrapartum fEEG. A normal baseline intrapartum fEEG activity was reported in several studies with evidence of alternance of sleep/wake states including two types of sleep behaviors (active and quiet) (34, 45, 49, 51, 60, 83). This "normal" intrapartum fEEG activity was similar to that of a newborn of the same age and same birth weight. Similarly, several studies identified patterns suggestive of fetal distress. Drugs, in particular, if given systemically, were shown to influence fEEG activity. Finally, a correlation was established between fEEG activity, FHR deceleration, Apgar scores (1 min and 5 min), and these factors were shown to be useful to predict the neurological outcome of the infants at 1 year of age. Animal studies using fetal sheep models and UCO were able to reproduce some of the abnormalities associated with fetal distress and showed that fEEG activity assessment could be a useful monitoring tool to help detect abnormal fetal brain activity associated with intrapartum complications.

The "normal" intrapartum fEEG activity was reported by several studies as a low voltage baseline pattern that varies from 5 to 50 μV per cm, with waves frequencies between 0.5 and 25 Hz. A predominant theta activity or an alternance of delta and theta activity were observed (45, 60). It is interesting to note that none of these early studies reported fEEG amplitude above 200 μV. We were able to record human intrapartum fEEG with a fetal scalp electrode with amplitudes around 400 μV. The data was acquired at 1,000 Hz. In this case, the amplitude of the raw signal is about twice the reported maximum of about 200 μV (45, 49, 83). It is possible that this high amplitude is the result of the effect of diazepam administration as reported

by Hopp et al. (30) and Khopp et al. (61). It is also possible that the older technologies and the filters used about 50 years ago might have prevented the capture of the intermittent faster waves with higher amplitude. This assumption is supported by the following: if we filter our recording similarly (i.e., 0.5–12 Hz), the fEEG tracing resembles more what these studies presented (amplitude below 200 μV) (**Figure 4**). A distinctive high-/low-frequency behavioral state pattern during the first stage of labor is seen as an alternance of 10 Hz and 2 Hz fEEG activity (**Figure 4**, TOP). It would, therefore, be interesting to repeat some of these older studies with the newest digital EEG technology.

Different studies have identified distinctive patterns suggestive of fetal distress and potentially associated with an abnormal outcome at 1 year of age. Particularly, sharp waves and long voltage depression were both reported to be more commonly identified in cases of fetal distress and neurologically abnormal children at 1 year of age (15, 23, 28, 29, 31).

The effect of intrapartum drug administration to the mother (for analgesia or anesthesia) was also reported in different studies and appeared more significant if the drug was given systemically (in comparison to local anesthesia) (26, 33, 34, 39, 40, 47).

One of the current limitations for routine use of fEEG monitoring remains the expertise required to read and interpret the tracing. Computer algorithms and methods to digitize the fEEG signal (including spectral analysis) have been developed but have remained experimental, failing to be translated to day-to-day practice (17, 32, 57). Computer-assisted fEEG reading and interpretation should be further developed to help democratize this tool allowing its routine use in an L&D unit.

The information gathered from fEEG, FHR monitoring, scalp pH measurements, Apgar score used as control measures of fetal health, and their relationships with one another were studied and the invaluable information they can provide have been demonstrated in several studies (15, 21, 23, 25, 26, 31). The 2019 Early Notification scheme progress report of the National Health System identified 70% of perinatal brain injuries as avoidable with continuous CTG monitoring (84). Therefore, both fEEG and FHR monitoring should ideally be part of the standard of care for intrapartum surveillance allowing earlier detection of fetal distress and identification of infants at risk of abnormal neurological long term outcomes to allow timely course corrections before the irreversible injury occurs. We emphasize the intentional choice of the term "ideally" in this recommendation, as the present reality in many L&D units is that continuous EFM is not utilized and scalp electrode placement is reserved to higher-risk deliveries. We suggest that this practice is to an extent the result of disagreement about the benefits of continuous EFM or scalp electrode placement and demonstration of such benefits for prevention of brain injuries may shift the preferences toward a broader adoption of these technologies.

Animal studies, and more precisely the ones using sheep model and UCO mimicking condition of fetal ischemia have proven useful to yield better knowledge of fEEG and its usefulness as a monitoring tool during labor (53, 54, 56, 57, 85). They are also useful in comparing treatment outcomes as shown by Gerrits et al. (55). However, while the benefit of translational medicine is indisputable, proper studies in human subjects and particularly prospective studies are still required to further establish the utility of fEEG monitoring intrapartum. Because this type of studies can be very challenging to conduct, in addition to the research aspects of fEEG, the research setting, and organization of the protocol are important for eventual success.

Recommendations for Successful Case Recruitment in Clinical Prospective Studies

Below we summarize our experience with conducting a prospective fEEG study at an L&D unit (**Figure 5**). The study recruitment process begins with two forms of passive engagement. A potential participant's first exposure to the study is an informational near the L&D reception desk. As the potential participant moves through the L&D ward, they will encounter bright purple door flyers denoting another occupant's participation in the study. Both of these engagements are low to medium impact and do not require interaction with

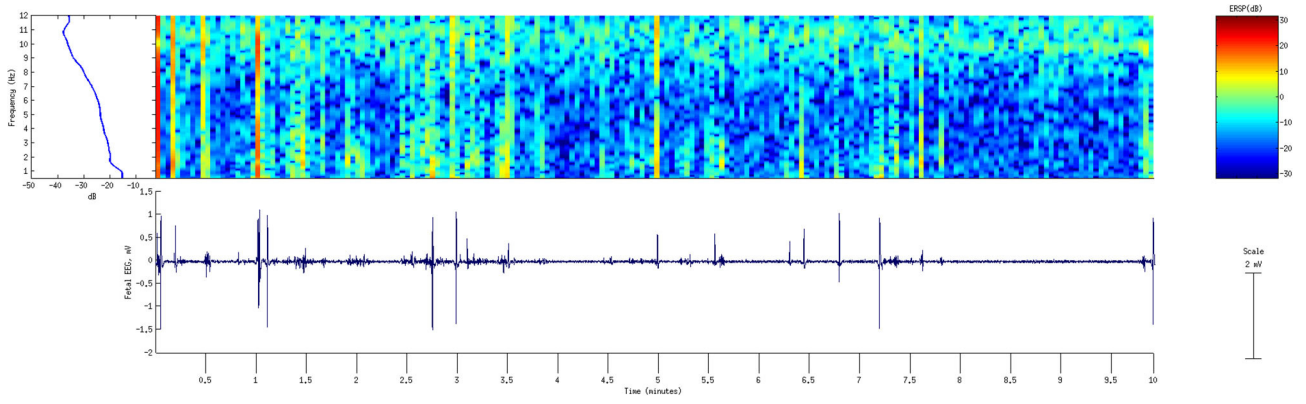

FIGURE 4 | FEEG recording from the standard fetal scalp electrode during the first stage of labor. A period of 10 min is shown with fEEG tracing (bottom) filtered 0.5–12 Hz and the corresponding power spectral analysis (top left) and wavelet transform (top left) to demonstrate the time-frequency behavior of fEEG. Note switching between delta and alpha-band activity. The X-axis shows time, with each segment corresponding to 0.5 min for a total of 10 min. Signal processing was performed in EEGLAB using Matlab 2013b, MathWorks, Mattick, MA. fEEG: fetal electroencephalogram.

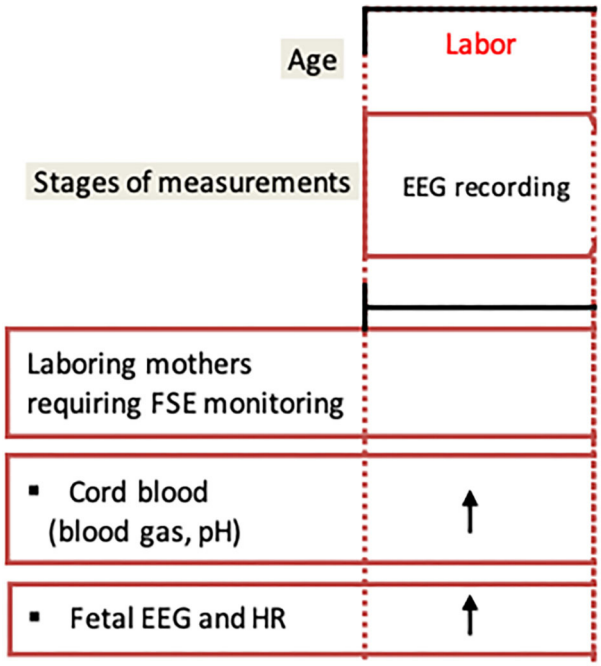

FIGURE 5 | Suggested study protocol. Fetal EEG recording during labor will be followed by cord blood measurements at birth to determine the degree of acidemia and the neonatal morbidity score. FSE, fetal scalp electrode; EEG, electroencephalogram; HR, heart rate.

study personnel. However, the name recognition and potential assurance of other families participating in the study lay the foundation for later direct interaction with study personnel. The next step in the recruitment process is this direct interaction. L&D staff identify potential study families and communicate the room numbers to recruiting personnel. This personnel approaches the family with an informed consent form and a summary sheet that further simplifies the objectives of the study. This step only gains initial interest from the family and is dependent on the placement of a fetal scalp electrode (FSE). If an FSE is utilized during the delivery, L&D staff will inform the technical study personnel to confirm consent and connect the study device. The device will record the data for future analysis.

Obtaining clinical data for assessing the automated algorithms proved more difficult than the initial study design anticipated. Recruitment of eligible families fell well below the initial study benchmarks. We, therefore, reviewed the consent and recruitment process to better communicate the goals of collecting the necessary data. Our review determined that simplifying a study's intervention down to the required effort by families and direct impact helped cut through many potential barriers to initial participant interest. Simple solutions included accompanying a three to five bullet-point summary sheet to complement the informed consent process. When reviewing the required informed consent form, the bullet point summary helped remind a laboring mother what the study required of her family. Another way we revised our process was to be cognizant of the laboring mother's attention span and the number of hospital personnel involved with the family's clinical visit. We retrained staff to keep interactions as brief as possible. Families are inundated with rounding clinical staff across multiple shifts; a lengthy interaction with study personnel for an optional study was likely to be dismissed by a laboring mother. It is also important to avoid approaching laboring women in the second stage of labor or after a significant deceleration when they are overwhelmed.

Upon review of other studies, we discovered that this was a common mistake in subject recruitment. Often, L&D studies overlook a subject's combination of being unfamiliar with their situation, being mentally/physically overwhelmed, and being unfamiliar with the consent process. We felt that our simple mitigating processes helped increase our potential subjects' interest.

Fetal EEG During Pregnancy

Although the focus of this review was on intrapartum fEEG studies, we cannot completely overlook the valuable information gathered from antepartum fetal magnetoencephalogram (fMEG) studies. As mentioned before, fEEG recording was first described by (14) in a 7-month fetus *in utero* and later described by Okamoto and Kirikae (86) who identified EEG activity in the fetus as early as 12 weeks old. In 1985, Blum et al. described a new technique, the fMEG, to record fetal brain electrical activity *in utero* (87). The technique had the benefit of being non-invasive, yet allowing to obtain fMEG traces of decent quality. The technique was further perfected by Eswaran et al. to minimize artifacts mainly from maternal ECG and fECG as well as from the environment (88). The fMEG allowed to study MEG patterns associated with fetal brain maturation similar to what is seen in preterm infants (89). The technique was also used to study behavioral states and sleep patterns associated with the gestational age allowing to gain better insight into the developing brain (90). These studies in humans are paralleled by findings in the fetal sheep model of human development where the relationship between fetal brain maturation and properties of fetal ECoG with regard to sleep states architecture and development was reported (91–93).

As the knowledge on antepartum fMEG/fEEG expands further, it will necessarily affect the more specific intrapartum situation. Therefore, to better understand intrapartum fEEG, staying up to date on the literature related to the antepartum EEG/MEG monitoring is necessary.

CONCLUSIONS

In this systematic review of the literature on intrapartum fEEG, we found that if a "normal" baseline EEG activity can be successfully recorded, abnormal patterns suggestive of fetal distress can also be observed. The combination of fEEG analysis with FHR monitoring as well as Apgar score can help identify patients at risk allowing early intervention. This should also help identify when the situation is not alarming, preventing unnecessary interventions such as C-section. The majority of the studies date back from the 70s with the potential that

some of their data could be invalidated by newer technologies. Despite the paucity of recent studies on the subject, the over 50 years of literature on fEEG clearly demonstrates that fEEG represents a clinically tested bedside monitoring technology of fetal well-being during labor with a clear potential to detect fetal distress, complementary to FHR monitoring. FEEG intrapartum warrants prospective clinical research with modern technical capabilities of data acquisition and computerized interpretation in L&D setting.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AC collected and screened all the studies, determined eligibility, summarized the findings from each study, compiled them in tables, translated the French language manuscripts, wrote, revised, and approved the final version of the manuscript. YF conducted the preliminary literature review by establishing the search method, keywords, started screening some of the studies, and contributed to the first draft and approved the final version of the manuscript. JF contributed to revised and approved the final version of the manuscript. FK contributed to reviewed and approved the final version of the manuscript. CA contributed to the manuscript, revised the manuscript, and approved its final version. MF conceived the manuscript, reviewed all studies screened, and confirmed if they were eligible or not and translated the German and Russian language manuscripts, drafted, revised, and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

MF's work on this subject has been funded, in part, by the Western Innovation Fund (WIF), Women's Development Council, London Health Sciences Centre, London, ON, Canada; Molly Towell Perinatal Research Foundation (MTPRF), Fonds de Recherche du Québec - Santé (FRQS), Canadian Institutes of Health Research (CIHR), the University of Washington's Dept. of Obstetrics and Gynecology. We gratefully acknowledge the Perinatal Research Lab of Dr. Bryan Richardson who had supported this research in the initial stages. The referenced fetal sheep UCO models and experiments came from this lab.

REFERENCES

1. Rees S, Inder T. Fetal and neonatal origins of altered brain development. *Early Hum Dev*. (2005) 81:753–61. doi: 10.1016/j.earlhumdev.2005.07.004
2. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet*. (2008) 371:261–9. doi: 10.1016/S0140-6736(08)60136-1
3. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*. (2018) 392:1789–858. doi: 10.1016/s0140-6736(18)32279-7
4. Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. *Lancet*. (2015) 385:430–40. doi: 10.1016/S0140-6736(14)61698-6
5. Parer JT, King T. Fetal heart rate monitoring: is it salvageable? *Am J Obstet Gynecol*. (2000) 182:982–7. doi: 10.1016/S0002-9378(00)70358-9
6. Low JA. Determining the contribution of asphyxia to brain damage in the neonate. *J Obstet Gynaecol Res*. (2004) 30:276–86. doi: 10.1111/j.1447-0756.2004.00194.x
7. Low JA, Killen H, Derrick EJ. Antepartum fetal asphyxia in the preterm pregnancy. *Am J Obstet Gynecol*. (2003) 188:461–5. doi: 10.1067/mob.2003.37
8. Shepherd E, Middleton P, Makrides M, McIntyre SJ, Badawi N, Crowther CA. Antenatal and intrapartum interventions for preventing cerebral palsy: an overview of Cochrane systematic reviews. *Cochrane Database Syst Rev*. (2016). doi: 10.1002/14651858.CD012077
9. 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) 216:163.e1–163.e6. doi: 10.1016/j.ajog.2016.10.009
10. Cahill AG, Mathur AM, Smyser CD, Mckinstry RC, Roehl KA, López JD, et al. Neurologic injury in acidemic term infants. *Am J Perinatol*. (2017) 34:668–75. doi: 10.1055/s-0036-1597135
11. Frasch MG. Saving the brain one heartbeat at a time. *J Physiol*. (2018) 596:5503–04. doi: 10.1113/JP275776
12. Gibbons L, Belizán JM, Lauer JA, Betrán AP. The global numbers and costs of additionally needed and unnecessary caesarean sections performed per year: overuse as a barrier to universal coverage. *World Health*. (2010). Available online at: https://www.researchgate.net/profile/Jose_Belizan/publication/265064468_The_Global_Numbers_and_Costs_of_Additionally_Needed_and_Unnecessary_Caesarean_Sections_Performed_per_Year_Overuse_as_a_Barrier_to_Universal_Coverage_HEALTH_SYSTEMS_FINANCING/links/549d75b30cf2fedbc31193f5/The-Global-Numbers-and-Costs-of-Additionally-Needed-and-Unnecessary-Caesarean-Sections-Performed-per-Year-Overuse-as-a-Barrier-to-Universal-Coverage-HEALTH-SYSTEMS-FINANCING.pdf
13. Spong CY, Berghella V, Wenstrom KD, Mercer BM, Saade GR. Preventing the first cesarean delivery: summary of a joint Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, and American College of Obstetricians and Gynecologists Workshop. *Obstet Gynecol*. (2012) 120:1181–93. doi: 10.1097/AOG.0b013e3182704880
14. Lindsley DB. Heart and brain potentials of human fetuses *in utero*. By Donald B. Lindsley, 1942. *Am J Psychol*. (1987) 100:641–6. doi: 10.2307/1422698
15. Rosen MG, Scibetta JJ, Hochberg CJ. Fetal electroencephalography. IV. The FEEG during spontaneous and forceps births. *Obstet Gynecol*. (1973) 42:283–9.
16. Eswaran H, Wilson JD, Lowery CL, Sharp G, Hawk RM, Murphy P, Pennington S. Brain stem auditory evoked potentials in the human fetus during labor. *Am J Obstet Gynecol*. (1999) 180:1422–6. doi: 10.1016/S0002-9378(99)70029-3
17. 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) 48:340–5. doi: 10.1203/00006450-200009000-00013
18. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, Moschetti I, et al. OCEBM levels of evidence working group. *The Oxford 2011 Levels of Evidence*. (2011).
19. Weller C, Dyson RJ, McFadyen IR, Green HL, Arias E. Fetal electroencephalography using a new, flexible electrode. *Br J Obstet Gynaecol*. (1981) 88:983–6. doi: 10.1111/j.1471-0528.1981.tb01685.x

20. Perinatologie/Geburtshilfe [Perinatology; obstetrics]. *Gynakol Rundsch*. (1981) 21 Suppl 3:55–78. doi: 10.1159/000269579
21. 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. *Br J Obstet Gynaecol*. (1979) 86:269–77. doi: 10.1111/j.1471-0528.1979.tb11254.x
22. Chik L, Sokol RJ, Rosen MG, Pillay SK, Jarrell SE. Trend analysis of intrapartum monitoring data: a basis for a computerized fetal monitor. *Clin Obstet Gynecol*. (1979) 22:665–79. doi: 10.1097/00003081-197909000-00013
23. Borgstedt AD, Heriot JT, Rosen MG, Lawrence RA, Sokol RJ. Fetal electroencephalography and one-minute and five-minute Apgar scores. *J Am Med Womens Assoc*. (1978) 33:220–2.
24. Nemeadze NO. Effect of the untimely bursting of waters on the functional state of the fetus in labor. *Akush Ginekol*. (1978) 7:44–8.
25. Chik L, Sokol RJ, Rosen MG, Regula GA, Borgstedt AD. Computer interpreted fetal monitoring data. Discriminant analysis or perinatal data as a model for prediction of neurologic status at one year of age. *J Pediatr*. (1977) 90:985–9. doi: 10.1016/S0022-3476(77)80577-5
26. Revol M, Challamel MJ, Fargier P, Bremond A, Chadenson O, Barrier PY. [An electroencephalographic study of the foetus during labour. Technique and interpretation (author's transl)]. *Rev Electroencephalogr Neurophysiol Clin*. (1977) 7:290–301. doi: 10.1016/S0370-4475(77)80008-7
27. Sokol RJ, Rosen MG, Chik L. Fetal electroencephalographic monitoring related to infant outcome. *Am J Obstet Gynecol*. (1977) 127:329–30. doi: 10.1016/0002-9378(77)90481-1
28. Chik L, Sokol RJ, Rosen MG, Borgstedt AD. Computer interpreted fetal electroencephalogram. I. Relative frequency of patterns. *Am J Obstet Gynecol*. (1976) 125:537–40. doi: 10.1016/0002-9378(76)90373-2
29. Chik L, Sokol RJ, Rosen MG, Borgstedt AD. Computer interpreted fetal electroencephalogram. II. Patterns in infants who were neurologically abnormal at 1 year of age. *Am J Obstet Gynecol*. (1976) 125:541–4. doi: 10.1016/0002-9378(76)90374-4
30. Hopp H, Beier R, Seidenschnur G, Heinrich J. [The significance of fetal electroencephalography for the diagnosis of fetal disorders]. *Zentralbl Gynakol*. (1976) 98:982–9.
31. Borgstedt AD, Rosen MG, Chik L, Sokol RJ, Bachelder L, Leo P. Fetal electroencephalography. Relationship to neonatal and one-year developmental neurological examinations in high-risk infants. *Am J Dis Child*. (1975) 129:35–8. doi: 10.1001/archpedi.1975.02120380021006
32. Chik L, Rosen MG, Sokol RJ. An interactive computer program for studying fetal electroencephalograms. *J Reprod Med*. (1975) 14:154–8.
33. Challamel MJ, Fargier P, Barrier PY, Revol M. [Fetal EEG during labor]. *Rev Electroencephalogr Neurophysiol Clin*. (1974) 4:429–33. doi: 10.1016/S0370-4475(74)80053-5
34. Fargier P, Bremond A, Challamel MJ, Barrier PY, Dolfus JM, Salomon B, Dargent D, Magnin P. L'électroencéphalogramme du fœtus au cours du travail et de l'accouchement. *J Gynéc Obstét Biol Repr*. (1974) 3:1023–33.
35. Heinrich J, Seidenschnur G. [Methods and results of intranatal intensive care. II. The RFT-fetal monitor BMT 504]. *Zentralbl Gynakol*. (1974) 96:513–23.
36. Beier R, Heinrich J, Hopp H, Seidenschnur G. [Problems in fetal electroencephalography]. *Psychiatr Neurol Med Psychol*. (1973) 25:92–7.
37. Carretti N, Arfoudi A, Gaja R. [Fetal electroencephalographic changes during labor after administration of oxygen to the mother]. *J Gynecol Obstet Biol Reprod*. (1973) 2:79–86.
38. Hopp H, Heinrich J, Seidenschnur G, Beier R. [Interpretation of simultaneously recorded fetal electroencephalograms and cardiotocograms]. *Zentralbl Gynakol*. (1973) 95:801–7.
39. Peltzman P, Goldstein PJ, Battagin R. Quantitative analysis of fetal electrophysiologic data. *Am J Obstet Gynecol*. (1973) 115:1117–24. doi: 10.1016/0002-9378(73)90562-0
40. Peltzman P, Goldstein PJ, Battagin R. Optical analysis of the fetal electroencephalogram. *Am J Obstet Gynecol*. (1973) 116:957–62. doi: 10.1016/S0002-9378(16)33843-1
41. Rosen MG, Scibetta JJ, Hochberg CJ. Fetal electroencephalography. IV. The FEEG during spontaneous and forceps births. *Obstet Gynecol*. (1973) 42:283–9.
42. 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) 115:37–47. doi: 10.1016/0002-9378(73)90086-0
43. Chachava KV, Devdariani MG, Zhordaniia ID, Loladze AS, Berulava AS. [Complex study of the mother and fetus in the 1st stage of physiological labor]. *Akush Ginekol*. (1972) 48:33–5.
44. Hopp H, Heinrich J, Seidenschnur G, Beier R, Schultz H. [Preliminary results of fetal electroencephalography and cardiotocography]. *Geburtshilfe Frauenheilkd*. (1972) 32:629–34.
45. Mann LI, Zwies A, Duchin S, Newman M. Human fetal electroencephalography: application of a vacuum electrode. *Am J Obstet Gynecol*. (1972) 114:898–903. doi: 10.1016/0002-9378(72)90094-4
46. Feldman JP, Le Houezec R, Sureau C. [Fetal electro-encephalography. Adjustment of an electrode allowing permanent recording during labor]. *Gynecol Obstet*. (1970) 69:491–3.
47. Rosen MG, Scibetta JJ, Hochberg CJ. Human fetal electroencephalogram. 3. Pattern changes in presence of fetal heart rate alterations and after use of maternal medications. *Obstet Gynecol*. (1970) 36:132–40.
48. Chachava KV, Devdariani MG, Loladze AS. [Several variants in the electroencephalogram of the fetus in normal and pathologic states]. *Akush Ginekol*. (1969) 45:18–21.
49. Rosen MG, Scibetta JJ. The human fetal electroencephalogram. I. An electrode for continuous recording during labor. *Am J Obstet Gynecol*. (1969) 104:1057–60. doi: 10.1016/0002-9378(69)90703-0
50. Barden TP, Peltzman P, Graham JT. Human fetal electroencephalographic response to intrauterine acoustic signals. *Am J Obstet Gynecol*. (1968) 100:1128–34. doi: 10.1016/S0002-9378(15)33414-1
51. Rosen MG, Satran R. Fetal electroencephalography during birth. *Obstet Gynecol*. (1965) 26:740–5.
52. De Haan HH, Gunn AJ, Williams CE, Gluckman PD. Brief repeated umbilical cord occlusions cause sustained cytotoxic cerebral edema and focal infarcts in near-term fetal lambs. *Pediatr Res*. (1997) 41:96–104. doi: 10.1203/00006450-199701000-00015
53. Thorngren-Jerneck K, Ley D, Hellström-Westas L, Hernandez-Andrade E, Lingman G, Ohlsson T, et al. Reduced postnatal cerebral glucose metabolism measured by PET after asphyxia in near term fetal lambs. *J Neurosci Res*. (2001) 66:844–50. doi: 10.1002/jnr.10051
54. Kaneko M, White S, Homan J, Richardson B. Cerebral blood flow and metabolism in relation to electrocortical activity with severe umbilical cord occlusion in the near-term ovine fetus. *Am J Obstet Gynecol*. (2003) 188:961–72. doi: 10.1067/mob.2003.219
55. Gerrits LC, Battin MR, Bennet L, Gonzalez H, Gunn AJ. Epileptiform activity during rewarming from moderate cerebral hypothermia in the near-term fetal sheep. *Pediatr Res*. (2005) 57:342–6. doi: 10.1203/01.PDR.0000150801.61188.5F
56. 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:e22100. doi: 10.1371/journal.pone.0022100
57. 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:e108119. doi: 10.1371/journal.pone.0108119
58. Kurz CS, Miltner FO. [Dynamics of the fetal EEG spectral pattern in normal delivery (author's transl)]. *Z Geburtshilfe Perinatol*. (1980) 184:401–12.
59. Bernstine RL, Borkowski WJ, Price AH. Prenatal fetal electroencephalography. *Am J Obstet Gynecol*. (1955) 70:623–30. doi: 10.1016/0002-9378(55)90357-4
60. Rosen MG, Scibetta JJ. The human fetal electroencephalogram. 2. Characterizing the EEG during labor. *Neuropadiatrie*. (1970) 2:17–26. doi: 10.1055/s-0028-1091837
61. Khopp K, Zaidenshnur G, Beier R, Kheinrikh I, Germann K. [Cardiotocographic and electroencephalographic studies in early diagnosis of fetal hypoxia]. *Akush Ginekol*. (1977) 10:27–31.
62. Infant Collaborative Group. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet*. (2017) 389:1719–29. doi: 10.1016/S0140-6736(17)30568-8

63. Huhmar E, Jaervinen PA. OBSERVATIONS ON FETAL ELECTROENCEPHALOGRAPHY. *Ann Chir Gynaecol Fenn*. (1963) 52:372–5.
64. Frasch MG, Durosier LD, Gold N, Cao M, Matushewski B, Keenliside L, et al. Adaptive shut-down of EEG activity predicts critical acidemia in the near-term ovine fetus. *Physiol Rep*. (2015) 3:e12435. doi: 10.14814/phy2.12435
65. Chik L, Sokol RJ, Rosen MG. Computer interpreted fetal electroencephalogram: sharp wave detection and classification of infants for one year neurological outcome. *Electroencephalogr Clin Neurophysiol*. (1977) 42:745–53. doi: 10.1016/0013-4694(77)90227-9
66. Morrison JL, Berry MJ, Botting KJ, Darby JRT, Frasch MG, Gatford KL, et al. Improving pregnancy outcomes in humans through studies in sheep. *Am J Physiol*. (2018) 315:R1123–53. doi: 10.1152/ajpregu.00391.2017
67. Back SA, Riddle A, Dean J, Hohimer AR. The instrumented fetal sheep as a model of cerebral white matter injury in the premature infant. *Neurotherapeutics*. (2012) 9:359–70. doi: 10.1007/s13311-012-0108-y
68. Pryds O. Control of cerebral circulation in the high-risk Neonate. *Ann Neurol*. (1991) 30:321–29. doi: 10.1002/ana.410300302
69. Menke J, Michel E, Hillebrand S, Von Twickel J. Cross-spectral analysis of cerebral autoregulation dynamics in high risk preterm infants during the perinatal Period1. *Pediatric*. (1997) doi: 10.1203/00006450-199711000-00023
70. du Plessis AJ. Cerebrovascular injury in premature infants: current understanding and challenges for future prevention. *Clin Perinatol*. (2008) 35:609–41. doi: 10.1016/j.clp.2008.07.010
71. Soul JS, Hammer PE, Tsuji M, Saul JP, Bassan H, Limperopoulos C, et al. Fluctuating pressure-passivity is common in the cerebral circulation of sick premature infants. *Pediatr Res*. (2007) 61:467–73. doi: 10.1203/pdr.0b013e31803237f6
72. Szymonowicz W, Walker AM, Yu VY, Stewart ML, Cannata J, Cussen L. Regional cerebral blood flow after hemorrhagic hypotension in the preterm, near-term, and newborn lamb. *Pediatr Res*. (1990) 28:361–6. doi: 10.1203/00006450-199010000-00012
73. Papile LA, Rudolph AM, Heymann MA. Autoregulation of cerebral blood flow in the preterm fetal lamb. *Pediatr Res*. (1985) 19:159–61. doi: 10.1203/00006450-198502000-00001
74. Tweed WA, Cote J, Pash M, Lou H. Arterial oxygenation determines autoregulation of cerebral blood flow in the fetal lamb. *Pediatr Res*. (1983) 17:246–9. doi: 10.1203/00006450-198304000-00002
75. Helou S, Koehler RC, Gleason CA, Jones MD Jr, Traystman RJ. Cerebrovascular autoregulation during fetal development in sheep. *Am J Physiol*. (1994) 266:H1069–74. doi: 10.1152/ajpheart.1994.266.3.H1069
76. Hohimer AR, Bissonnette JM. Effects of cephalic hypotension, hypertension, and barbiturates on fetal cerebral blood flow and metabolism. *Am J Obstet Gynecol*. (1989) 161:1344–51. doi: 10.1016/0002-9378(89)90695-9
77. Dhillon SK, Lear CA, Galinsky R, Wassink G, Davidson JO, Juul S, et al. The fetus at the tipping point: modifying the outcome of fetal asphyxia. *J Physiol*. (2018) 596:5571–92. doi: 10.1113/JP274949
78. Gunn AJ, Maxwell L, De Haan HH, Bennet L, Williams CE, Gluckman PD, et al. Delayed hypotension and subendocardial injury after repeated umbilical cord occlusion in near-term fetal lambs. *Am J Obstet Gynecol*. (2000) 183:1564–72. doi: 10.1067/mob.2000.108084
79. *Safe Prevention of the Primary Cesarean Delivery - ACOG*. Available online at: https://www.acog.org/Clinical-Guidance-and-Publications/Obstetric-Care-Consensus-Series/Safe-Prevention-of-the-Primary-Cesarean-Delivery (accessed May 24, 2018).
80. Yumoto Y, Satoh S, Fujita Y, Koga T, Kinukawa N, Nakano H. Noninvasive measurement of isovolumetric contraction time during hypoxemia and acidemia: fetal lamb validation as an index of cardiac contractility. *Early Hum Dev*. (2005) 81:635–42. doi: 10.1016/j.earlhumdev.2005.04.004
81. Astrup J. Energy-requiring cell functions in the ischemic brain. Their critical supply and possible inhibition in protective therapy. *J Neurosurg*. (1982) 56:482–97. doi: 10.3171/jns.1982.56.4.0482
82. Attwell D, Laughlin SB. An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab*. (2001) 21:1133–45. doi: 10.1097/00004647-200110000-00001
83. Rosen MG, Scibetta J. Documenting the human fetal EEG during birth. *Electroencephalogr Clin Neurophysiol*. (1969) 27:661. doi: 10.1016/0013-4694(69)91233-4
84. Resolution NHS. *The Early Notification Scheme Progress Report: Collaboration and Improved Experience for Families*. Sep 2019. Available online at: https://resolution.nhs.uk/wp-content/uploads/2019/09/NHS-Resolution-Early-Notification-report.pdf
85. Paradise JL, Campbell TF, Dollaghan CA, Feldman HM, Bernard BS, Colborn DK, et al. Receptive vocabulary and cognition in 3-year-old children in relation to otitis media in their first 3 years of life. *Pediatr Res*. (1997) 41:96. doi: 10.1203/00006450-199704001-00584
86. Okamoto Y, Kirikae. Electroencephalographic studies on brain of foetus, of children of premature birth and new-born, together with a note on reactions of foetus brain upon drugs. *Folia Psychiatr Neurol Jpn*. (1951) 5:135–46. doi: 10.1111/j.1440-1819.1951.tb00583.x
87. Blum T, Saling E, Bauer R. First magnetoencephalographic recordings of the brain activity of a human fetus. *Br J Obstet Gynaecol*. (1985) 92:1224–9. doi: 10.1111/j.1471-0528.1985.tb04866.x
88. Preissl H, Lowery CL, Eswaran H. Fetal magnetoencephalography: current progress and trends. *Exp Neurol*. (2004) 190 Suppl 1:S28–36. doi: 10.1016/j.expneurol.2004.06.016
89. Eswaran H, Haddad NI, Shihabuddin BS, Preissl H, Siegel ER, Murphy P, et al. Non-invasive detection and identification of brain activity patterns in the developing fetus. *Clin Neurophysiol*. (2007) 118:1940–6. doi: 10.1016/j.clinph.2007.05.072
90. Haddad N, Govindan RB, Vairavan S, Siegel E, Temple J, Preissl H, et al. Correlation between fetal brain activity patterns and behavioral states: an exploratory fetal magnetoencephalography study. *Exp Neurol*. (2011) 228:200–5. doi: 10.1016/j.expneurol.2011.01.003
91. Keen AE, Frasch MG, Sheehan MA, Matushewski B, Richardson BS. Maturational changes and effects of chronic hypoxemia on electrocortical activity in the ovine fetus. *Brain Res*. (2011) 1402:38–45. doi: 10.1016/j.brainres.2011.05.043
92. Keen AE, Frasch MG, Sheehan MA, Matushewski BJ, Richardson BS. Electrocortical activity in the near-term ovine fetus: automated analysis using amplitude frequency components. *Brain Res*. (2011) 1402:30–7. doi: 10.1016/j.brainres.2011.05.044
93. Rao N, Keen A, Czikk M, Frasch M, Richardson BS. Behavioural state linkage in the ovine fetus near term. *Brain Res*. (2009) 1250:149–56. doi: 10.1016/j.brainres.2008.11.003

Conflict of Interest: MF holds a patent "EEG Monitor of Fetal Health" US9,215,999 and has a start-up Health Stream Analytics LLC to develop fetal EEG technologies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 09 December 2020
doi: 10.3389/fped.2020.599049

Non-invasive Fetal Electrocardiography for Intrapartum Cardiotocography

Rik Vullings[1,2*] *and Judith O. E. H. van Laar*[3]

[1] Biomedical Diagnostics Lab Eindhoven, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands, [2] Nemo Healthcare, Veldhoven, Netherlands, [3] Máxima Medical Center, Veldhoven, Netherlands

Fetal monitoring is important to diagnose complications that can occur during pregnancy. If detected timely, these complications might be resolved before they lead to irreversible damage. Current fetal monitoring mainly relies on cardiotocography, the simultaneous registration of fetal heart rate and uterine activity. Unfortunately, the technology to obtain the cardiotocogram has limitations. In current clinical practice the fetal heart rate is obtained via either an invasive scalp electrode, that poses risks and can only be applied during labor and after rupture of the fetal membranes, or via non-invasive Doppler ultrasound technology that is inaccurate and suffers from loss of signal, in particular in women with high body mass, during motion, or in preterm pregnancies. In this study, transabdominal electrophysiological measurements are exploited to provide fetal heart rate non-invasively and in a more reliable manner than Doppler ultrasound. The performance of the fetal heart rate detection is determined by comparing the fetal heart rate to that obtained with an invasive scalp electrode during intrapartum monitoring. The performance is gauged by comparing it to performances mentioned in literature on Doppler ultrasound and on two commercially-available devices that are also based on transabdominal fetal electrocardiography.

Keywords: electrophysiology, cardiotocography, fetal heart rate, fetal electrocardiogram, signal processing, artificial intelligence

OPEN ACCESS

Edited by:
Ines Maria Nunes, North Maternal Child Center Albino Aroso (CMIN), Portugal

Reviewed by:
Hernâni Gonçalves, University of Porto, Portugal
Vikranth Bapu Anna Venugopalan, Sandwell & West Birmingham Hospitals NHS Trust, United Kingdom

***Correspondence:**
Rik Vullings
r.vullings@tue.nl

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *26 August 2020*
Accepted: *19 November 2020*
Published: *09 December 2020*

Citation:
Vullings R and van Laar JOEH (2020) Non-invasive Fetal Electrocardiography for Intrapartum Cardiotocography. Front. Pediatr. 8:599049. doi: 10.3389/fped.2020.599049

1. INTRODUCTION

One in every five pregnant women experiences complications during her pregnancy (1). Although most of these complications are relatively harmless, some are more severe and will lead to fetal morbidity, or even mortality. The most important pregnancy complications, in terms of severity and occurrence, are premature birth, birth hypoxia, intrauterine growth restriction, and congenital anomalies. Together, this "big four" of pregnancy complications accounts for the majority of perinatal morbidities and mortalities (2).

Early detection of these pregnancy complications is of the utmost importance to prevent irreversible damage, but is unfortunately hampered by limitations of the technology that is used in daily clinical practice. Essentially, this technology comprises of cardiotocography and ultrasound imaging. The former constitutes a simultaneous registration of fetal heartrate (FHR) and maternal uterine activity (UA). It is used to screen for patterns in FHR or heartrate variability that could reveal a compromised condition, e.g., acidaemia (3). The latter is mostly used to screen for anomalies such as growth restriction or congenital heart disease (CHD).

The cardiotocogram (CTG) is obtained in daily practice by either invasive means, using a fetal scalp electrode (FSE) and intrauterine pressure catheter (IUPC), or non-invasive means, using a Doppler ultrasound probe and an external tocodynamometer. The invasive methods suffer from the limitations that they can only be used during labor after rupture of the fetal membranes and impose risks to mother and child (4). In some countries, these invasive methods are therefore no longer used. The non-invasive methods can be used throughout pregnancy, but are known to be unreliable (5, 6).

The Doppler ultrasound employed in cardiotocography consists of a rather narrow beam of ultrasound that insonifies a small volume in the maternal abdomen (7). If the fetal heart is within this volume, the signal to noise ratio of the reflected ultrasound beam is typically good enough to extract a reliable FHR. However, movement of mother or fetus, or high body mass of mother causes poor insonification of the fetal heart, with corresponding poor reliability of the derived FHR (8, 9). Also in preterm fetuses or multiple pregnancies, Doppler ultrasound is known to perform poorly.

Over the past decades, extensive research has focused on non-invasive fetal electrophysiological recordings for measurement of the CTG (10). Other than Doppler ultrasound, these electrophysiological recordings are hardly affected by movement and body mass (11, 12). However, the recordings are corrupted by many electrical interferences, of which the maternal heart is the dominant source. Many studies have been published on methods to remove this interference, i.e., the maternal electrocardiogram (ECG), and virtually all with good performance (13–18). Yet, in many practical situations, removal of the maternal ECG alone is not enough to enable reliable measurement of the FHR (18, 19). For example, during labor the maternal abdominal muscles cause interferences that exceed the fetal electrical cardiac activity (i.e., fetal ECG) in terms of amplitude and that overlap in the frequency domain.

Perhaps due to these practical limitations, to date, only a few solutions exist that constitute an electrophysiology-based device for CTG acquisition and that are ready for use in clinical practice. A few examples of such solutions include the GE Novii (GE, USA, formerly the Monica Healthcare Novii), the Philips Avalon Beltless (Philips, the Netherlands), and the Nemo Fetal Monitoring System (Nemo Healthcare, the Netherlands, of which one of the authors is co-founder). The use of these solutions in clinical practice is still fairly limited, mainly due to the relatively poor performance during second stage of labor (6, 12).

This paper proposes a practical solution to solve the problems that limit the application of non-invasive fetal electrocardiography-based cardiotocography, especially during second stage of labor. The focus of the paper lies on the acquisition of FHR; for assessing the maternal uterine activity, the reader is referred to literature such as (20–22). The performance of the method is assessed by comparing the FHR to that determined with a simultaneously applied FSE during intrapartum monitoring.

This paper is organized as follows: In section 2, the methodology for acquiring electrophysiological data and the signal processing toward cardiotocography are discussed and details are provided on the datasets used in this paper. In section 3, results of the signal processing methods are illustrated and in section 4 the results are discussed.

2. MATERIALS AND METHODS

The various methodological steps that are needed for acquiring electrophysiological data and signal processing toward FHR are schematically depicted in **Figure 1**. These steps will be discussed in more detail in the subsections below. Because some of these steps have been described in detail in other publications, we will discuss the methodology in terms of a "cookbook recipe" and will focus our description on the steps that have not yet been published in detail, i.e., the FHR detection step with artificial intelligence (AI) extension.

2.1. Data

The study protocol for the data used in this study was approved by the institutional review board of the Máxima Medical Center in December 2017 (NL63732.015.17). Women in established labor, carrying a healthy singleton fetus in cephalic presentation and with a gestational age between 36 and 42 weeks were eligible to participate. After written informed consents, participants received an adhesive electrode patch (Nemo Healthcare BV, the Netherlands) that comprises four unipolar electrodes, a ground electrode, and a common reference. Data were recorded locally on the patch and digitized at 500 Hz sample rate with a resolution

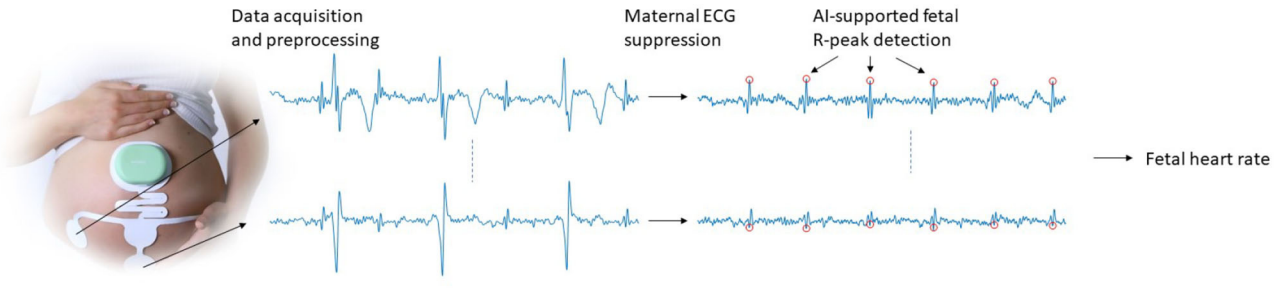

FIGURE 1 | Schematic illustration of the data acquisition and signal processing steps that are used to obtain FHR.

of 22 nV. Subsequently, data was transmitted wirelessly to a data processing device to yield instantaneous output of FHR, maternal heart rate, and UA. In parallel, digitized signals were stored on the data acquisition system to enable offline processing. All results in this paper are obtained via offline processing of stored data to allow for quantitative evaluation. To facilitate such evaluation, for all patients, a simultaneous FHR recording using a FSE was performed. The FHR of the scalp electrode was determined by a Philips Avalon FM30 cardiotocograph (Philips, the Netherlands) at sampling period of 4 Hz. The output of the cardiotocograph was stored digitally as well.

In total, 136 recordings were performed with an average duration of 185 ± 135 min, with the shortest recording 17 min and the longest 600 min. The evaluation of the presented method was only performed on 26 of these 136 recordings. The other 110 measurements were used to develop the methods presented in this study. More details on this splitting of the dataset are provided in section 2.2.3.2. Details on the age and body-mass-index of the mother are provided in **Table 1**; other relevant details such as fetal gender, weight, and presentation were unfortunately not registered.

2.2. Signal Processing

2.2.1. Preprocessing

The recorded signals were preprocessed to suppress interferences from e.g., abdominal muscles, baseline wander, and mains powerline. This preprocessing consists of the application of zero-phase delay highpass and lowpass filters, with cutoff frequencies of 1 and 70 Hz, respectively.

For the mains powerline, depending on geographical location (e.g., 50 Hz for Europe, 60 Hz for USA), a Kalman smoother was used to effectively suppress the powerline interference, while avoiding so-called ringing that characterizes conventional (in)finite impulse response filters (23). Details of the applied Kalman smoother are provided in (23).

2.2.2. Maternal ECG Suppression

After preprocessing, the dominant interference in the electrophysiological abdominal recordings is the maternal ECG. As mentioned in section 1, many studies have been published on methods for suppressing the maternal ECG. Most of these methods perform good enough to the point where the (possible) residuals of maternal ECG are no longer the dominant interference and where the methods do not cause any significant degradation to the quality of the remaining fetal ECG.

In this work, we use a template-based maternal ECG suppression method. First maternal QRS complexes are detected using a low-complexity R-peak detection method, presented in (24). Then the recorded signals are segmented, based on the detected maternal R-peaks, to yield one maternal ECG complex per segment. Each ECG complex is then further segmented to yield individual ECG waves. For each wave, a template is generated from the linear prediction of corresponding waves from preceding ECG complexes. The wave templates are subsequently combined to yield a template ECG. This method is discussed in detail in (15). Because the FHR is typically not correlated to the maternal heart rate, fetal ECG complexes occur in random places in the maternal ECG segments. In the linear prediction step, these fetal ECG complexes are therefore strongly attenuated in the template.

As a final step in the maternal ECG suppression, the templates per ECG complex are concatenated to produce an estimate of the maternal ECG signal which is then subtracted from the recorded signal, ideally preserving the fetal ECG. This procedure is illustrated in **Figure 2**.

2.2.3. Fetal Heart Rate Detection

Despite the accurate maternal ECG suppression that can be achieved, often the fetal ECG is still obscured by other interferences that remain after maternal ECG cancelation. In such cases, reliable detection of the FHR is still challenging. In **Figure 3** and example of a low-quality fetal ECG signal is shown.

2.2.3.1. Hierarchical Probabilistic Framework for R-peak Detection

In (19), we have introduced a model-based approach for detecting the fetal R-peaks. This approach leverages models on the fetal QRS waveform, on the heartrate, and on the noise dynamics to yield a robust fetal R-peak detection, even in case of low-quality signals. In this work, we extend our previous method with AI to further improve its robustness.

The method by Warmerdam et al. is based on the following state-space equation:

$$\mu_{k+1} = \mu_k + \vec{w}_k\vec{\theta}_{k+1} + \nu_{k+1} \quad (1)$$

$$\vec{y}_{k+1} = G\left(\vec{t}, \mu_{k+1}, \vec{z}_{k+1}\right) + \vec{\xi}_{k+1}. \quad (2)$$

In Equation (1), μ_{k+1} is the location of the $(k+1)^{\text{th}}$ fetal R-peak, $\vec{w}_k$ are previously detected interbeat (i.e., RR) intervals, $\vec{\theta}_{k+1}$ are the coefficients from an autoregressive (AR) model, and ν_{k+1} is a term that accounts for heartrate variability. Essentially, the location of the next fetal R-peak is estimated to be the location of the previous R-peak, plus the expected RR-interval, plus a random term. This random term ν_{k+1} is sampled from a zero-mean normal distribution.

In Equation (2), $\vec{y}_{k+1}$ is the $(k+1)^{\text{th}}$ segment of the recorded signal y, $G(\cdot)$ is a function that describes the linear combination of three Gaussian functions:

TABLE 1 | Age and body-mass-index (BMI) of the patients included in the study, subdivided over patients that were used in the training of the proposed methods and patients that were used in the evaluation of the methods.

	n	Age (years)	BMI ($kg \cdot m^{-2}$)	*p*-value
Training	110*	31.2 ± 4.4	28.5 ± 5.0	0.74
Testing	26†	30.9 ± 2.3	28.1 ± 5.5	0.70

The number of patients per group is indicated by n. A unpaired t-test was performed to determine whether the age and BMI between train and test datasets were significantly different.

**For two subjects from the train dataset the age was not known, for one subject the BMI was not known.*

†For one subject from the test dataset the BMI was not known.

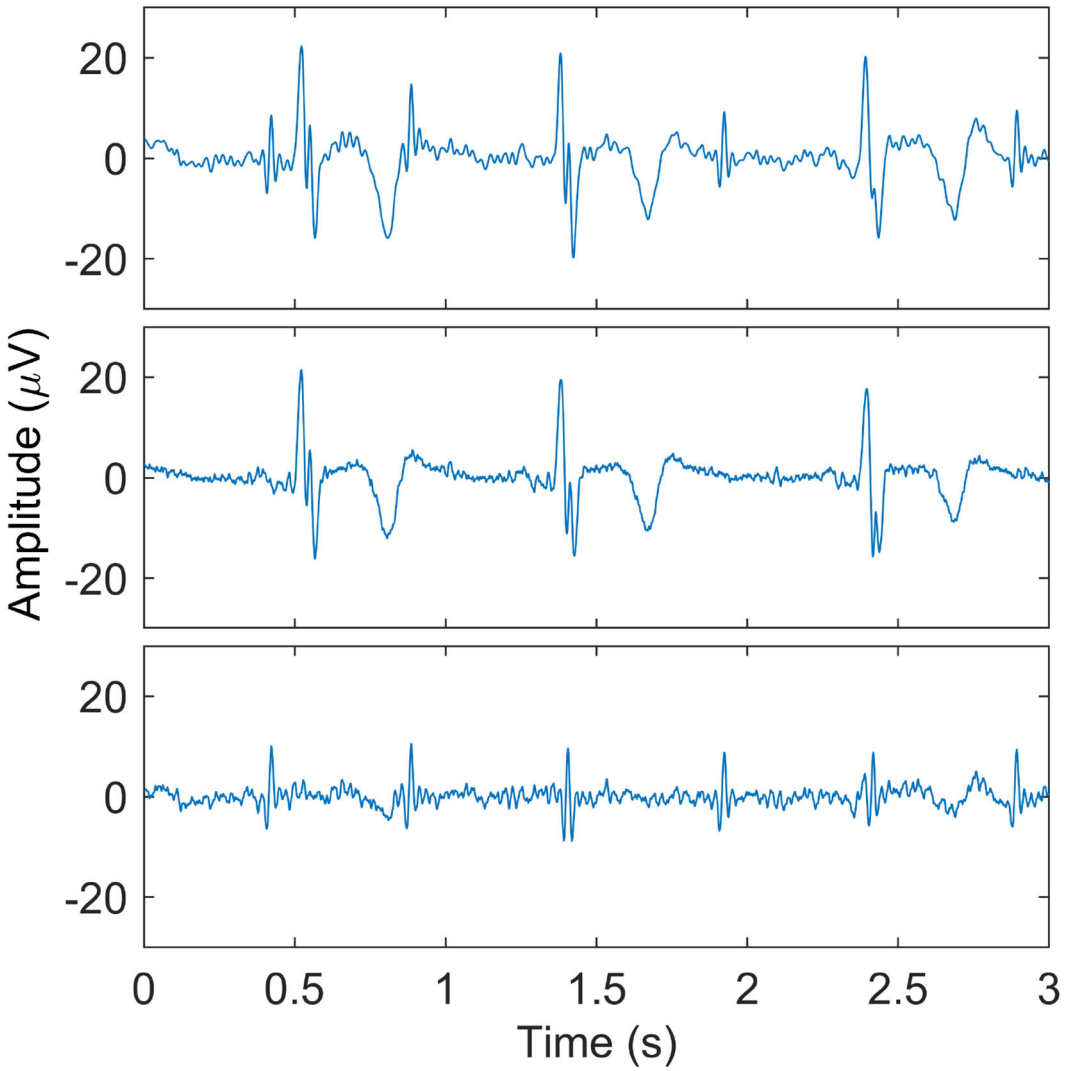

FIGURE 2 | Illustration of maternal ECG estimation. In the **(top)** panel, one of the recorded signals after preprocessing is shown. The **(middle)** panel shows the estimate of the maternal ECG signal, and the **(bottom)** panel shows the signal that results after subtracting the maternal ECG estimate. Here, the fetal QRS complexes are clearly visible.

$$G(t,\mu_k,\vec{z}) = \left(a_1 + a_2\,(t-\mu_k) + a_3\left(1 - \frac{(t-\mu_k)^2}{b^2}\right)\right) e^{-\frac{(t-\mu_k)^2}{2b^2}}, \tag{3}$$

$\vec{z}_{k+1} = [a_1, a_2, a_3, b]$ are the parameters (i.e., amplitudes $\vec{a}$ and variance b) for these Gaussian functions, and $\vec{\xi}_{k+1}$ represents measurement noise.

Using Bayes' rule, the *maximum a posteriori* estimate $\hat{\mu}_{k+1}$ can be obtained as:

$$\hat{\mu}_{k+1} = \arg\max_{\mu} \left\{ -\frac{\left(\mu_{k+1} - \hat{\mu}_k + \hat{w}_{k+1}^T \vec{\theta}_{k+1}\right)^2}{\Gamma^{\mathrm{HR}}} - \left(\vec{y}_{k+1} - \hat{\vec{y}}_{k+1}\right)^T \Gamma^{\mathrm{QRS}^{-1}} \left(\vec{y}_{k+1} - \hat{\vec{y}}_{k+1}\right) \right\}. \tag{4}$$

Here, Γ^{HR} and Γ^{QRS} are estimates of the heart rate variability and measurement noise, respectively, and $\hat{\vec{y}}_{k+1}$ is the estimate of the recorded signal (cf. Equation 2). The first term on the righthandside is in following paragraphs referred to as the *prior* model and the second term the *likelihood* model. All model parameters are updated using (extended) Kalman filters. For exact details on this method, the reader is referred to (19).

2.2.3.2. Artificial Intelligence Extension

Although the method by Warmerdam et al. was designed and shown to be robust against low-quality signals, situations occur where its performance rapidly decreases. This happens for instance when noise or interferences cause erroneous updates of the model parameters. At that point, a vicious circle will cause the next R-peak detection to go wrong, which in turn further diverges the model parameters, and so on. Therefore, in

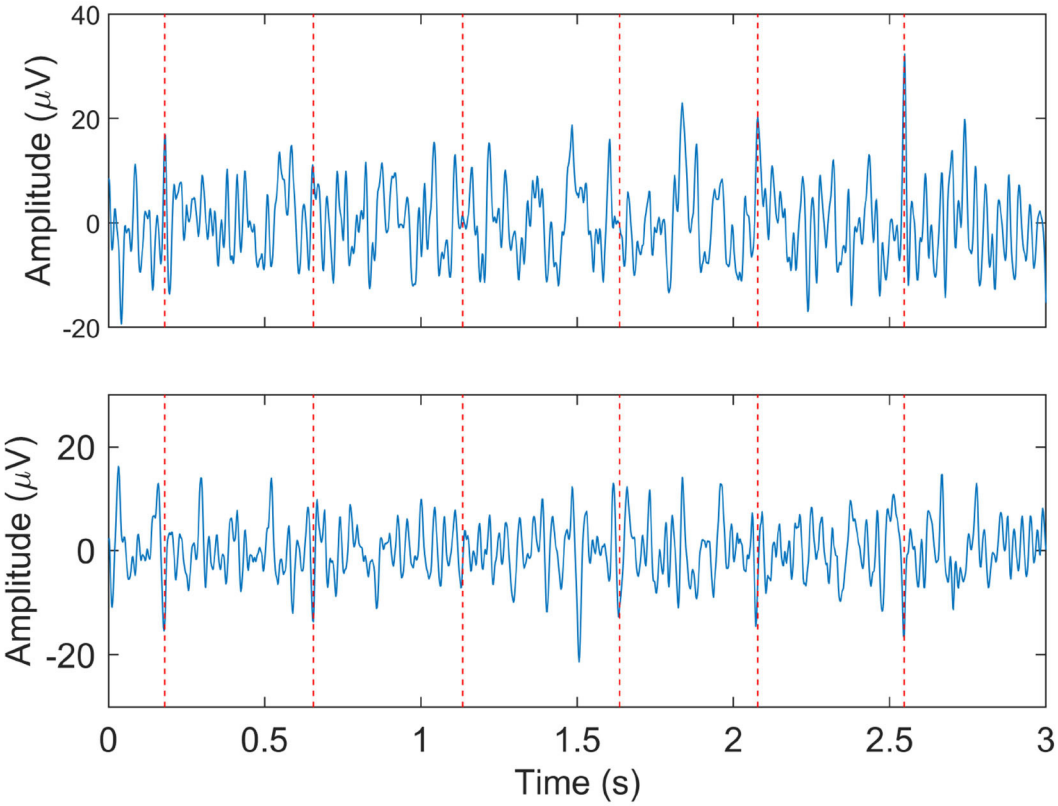

FIGURE 3 | Example of relatively low quality fetal ECG signals after maternal ECG suppression. The two panels show the signals from two different electrodes at the same moment. The dashed vertical lines indicate the locations at which the fetal heart rate detection method did detect fetal R-peaks.

this paper we propose an AI-based extension of the method to prevent such scenarios.

In this AI-based extension the RR-intervals are estimated using an AI model that is described in (25). Although this model at first seems to outperform other methods, its main limitation is that it can provide FHR outputs that look physiologically plausible but are in fact incorrect (25). Yet, in this work we use the RR-intervals detected by the AI model to validate the RR-interval estimate by the AR model: $\hat{w}_{k+1}^T \vec{\theta}_{k+1}$. As described above, heartrate variability is modeled in the state-space representation as random values v_{k+1} sampled from a zero-mean normal distribution. Assuming that this distribution has variance Σ_{k+1}, in case of poor agreement between the AR and AI model, this variance is increased as:

$$\Sigma_{k+1} \leftarrow \Sigma_{k+1} + \left| \hat{w}_{k+1}^T \vec{\theta}_{k+1} - RR_{\mathrm{AI}_{k+1}} \right|, \qquad (5)$$

where RR_{AI} is the RR-interval determined by the AI model. In case the two models are not in agreement, the variance of the prior model in Equation (4) is increased to the degree where this distribution can become virtually flat and no prior knowledge on the location of the new fetal R-peak is assumed. New fetal R-peak detections will therefore only be based on the agreement between the expected shape $\hat{\vec{y}}_{k+1}$ of the fetal R-peak and the recorded signal $\vec{y}_{k+1}$.

To prevent erroneous updates of the model parameters, when the difference between the RR-intervals detected by the presented method and the AI model exceeds 0.05 s, the model parameters are not updated. Likewise, no FHR output is shown to the clinician to prevent showing unreliable FHR information.

To train the AI model, 110 of the 136 recordings that had simultaneous FHR recording with the presented method as well as with FSE were randomly selected. The remaining 26 recordings were used as holdout set to evaluate the performance. A validation of only 26 recordings is relatively small, albeit that these 26 recordings together comprise of 84.2 h of multi-channel abdominal fetal electrophysiological recordings. With common use of such AI, a significant risk of overfitting to the training data might occur. In the method proposed here, this risk is largely mitigated by using the RR-intervals that are determined by the AI to increase the variance Σ_{k+1}. In case the RR-intervals would be overfitted, this variance would increase and the fetal R-peak will be based more on the *likelihood* model of Equation (4). Yet, to provide insights in the potential overfitting of the AI to the training data, in the section 3 we will provide results from the validation data as well as from the training data.

2.2.3.3. Postprocessing of Fetal Heart Rates

The FHR can be calculated from the detected fetal R-peak positions, yielding a FHR on a beat-to-beat basis. However, to

facilitate the communication with central monitoring systems (CMS) and ease the comparison to other methods for which the FHR has been acquired via CMS, this beat-to-beat FHR was resampled to 4 Hz using linear interpolation. Prior to resampling, outliers, which were defined as FHR values that differ more than 20 % from the previous FHR values (26), were omitted and replaced by zeros.

2.3. Methodology for Evaluation

For evaluation of the presented method, the performance of FHR detection can be assessed by comparing the FHR to that of the FSE. Moreover, the performance of our method can be gauged by comparing it to that of other methods reported in literature. However, the various models in our method are initialized such that they work optimally when always the same electrode positions are chosen. This is also illustrated in **Figure 1**, where it is shown that a single electrode patch is used to guarantee consistent placement of the electrodes. Because of this limitation, we cannot apply our method to public datasets such as the Physionet Non-invasive Fetal ECG Database, as in this dataset "electrode positioning was varied in order to improve SNR" (27). Yet, we can compare our method to results from (6) and (12) where similar devices are tested on similar datasets. Both devices from these studies (i.e., Monica Healthcare AN24 and Nemo Fetal Monitoring System) use transabdominal electrodes to record a multi-channel fetal ECG and use proprietary signal processing methods to extract the FHR from these recordings. In fact, with respect to the study of Lempersz et al., our work presents an extension of the algorithms and datasets presented in that paper. With respect to the study of Cohen et al., it should be noted that the comparison on FHR detection performance is only indirect because different datasets are used.

As evaluation metrics, we opt to express the performance in FHR estimation in terms of *success rate*, *reliability*, and *accuracy*. Here, *success rate* is defined as the percentage of time the method can provide a FHR estimation (6, 12). *Reliability* is expressed in terms of *positive percent agreement (PPA)* which is defined as the percentage of FHR values provided by the method that are within a 10% margin from a valid simultaneous FHR from the FSE (6). For *accuracy* we use bootstrapping of the absolute differences between the FHR from our method and that of the FSE. This metric is different from the definitions by Cohen et al. and by Lempersz et al., which are also different from each other. The reason for choosing a different way of calculating the *accuracy* is described below.

In Cohen et al., the *accuracy* is determined by the root-mean-squared error of the difference between the FHR from two devices vs. the expected difference that is determined by regression in the Bland-Altman plot. In case of a bias between the two FHR measurements, the regression in the Bland-Altman plot will correct for this, yielding a very small metric, even when the FHR determined from the non-invasive measurements differs significantly from the FHR from the FSE. In their paper, Cohen et al. did present the slope and y-intercept of the regression plots, making it possible to appropriately appreciate their findings, but in this work we prefer to show the *accuracy* as a metric that can be interpreted independent from other metrics such as the slope of the regression.

In Lempersz et al. the *accuracy* is determined based on bootstrapping over differences between the FHR from the non-invasive measurements and that of the FSE. If the FHR from the non-invasive measurements would be inaccurate, but without a significant bias, again the metric would be very small. By using the absolute difference instead of the signed differences, this issue is resolved.

Next to mean and standard deviations of the *success rate*, *reliability*, and *accuracy*, we also provide 95% confidence intervals (CI) and for the *accuracy* limits of agreements. For *accuracy*, all analyses are done using bootstrapping (28). Each bootstrap sample was generated by drawing a random pair of non-invasive FHR and FSE FHR for each woman included in the analysis. For the bootstrap sample, mean absolute difference and standard deviation of absolute differences were determined. This process was repeated 10,000 times to yield a large distribution for the mean absolute difference. The average accuracy and 95% CI were determined by taking the mean of the distribution and the 2.5 and 97.5% centile of the distribution, respectively. From the 10,000 bootstrap samples, also the mean standard deviation was determined, which was subsequently used to calculate limits of agreement as mean accuracy $\pm$ 2 $\times$ mean standard deviation (12).

3. RESULTS

In **Figures 4**, **5**, two examples of FHR tracings that were obtained with the presented method are shown relative to the FHR that was simultaneously obtained with a FSE. These two examples are from different patients and show FHR during the first and second stage of labor, respectively.

It can be seen in both these Figures that the resemblance between FHR patterns obtained with the presented method is high compared to the FHR patterns obtained with FSE. When looking in more detail, in **Figure 5** it can be seen that the FHR deceleration at 67.5 min is slightly underestimated by the presented method. While the FSE reveals a drop in FHR to 65 beats-per-minute (bpm), the presented methods shows a drop to 75 bpm. Despite this difference, the depicted FHR patterns can be considered to be clinically equivalent.

In **Tables 2–4**, the results of the quantitative comparisons between the developed method and the reference methods are provided. In **Table 2**, the overall results are provided, while in **Tables 3**, **4** the results are divided in first and second stage of labor, respectively.

As mentioned before, the tables do not only provide the results of the presented method on the validation set but also on the train set, to enable the assessment of potential overfitting of the AI to the train set. When comparing the results for the train and validation set, it can be argued that these are comparable and hence the risk that the results are indeed overfitted is small. In fact, the performance on the validation set might be even slightly better than that on the train set.

4. DISCUSSION

In this paper a new modular methodology for non-invasive electrophysiology for FHR acquisition was described. The method consists of various modules that have been individually developed and published, but with the ultimate goal of reliable FHR monitoring in mind. In the current paper, the mutual dependencies of the modules are described and an improved module for FHR detection was described that, based on the results, makes a relatively large difference in performance.

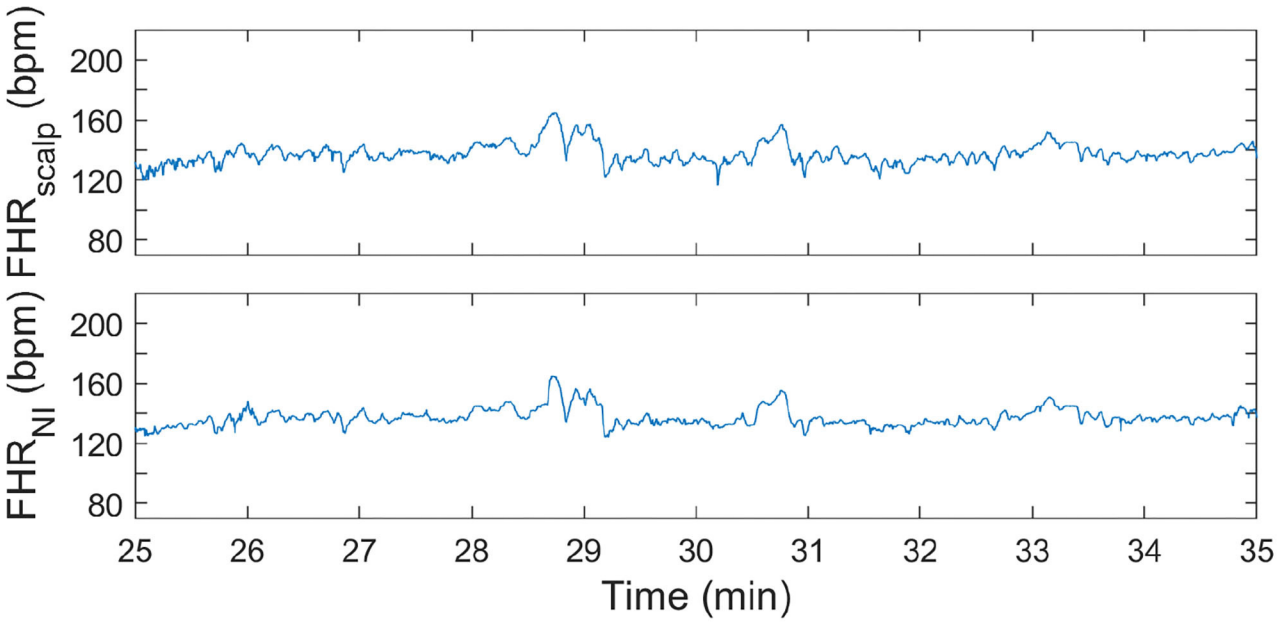

FIGURE 4 | FHR tracing during first stage of labor. In the **(top)** panel, the FHR from FSE is depicted. In the **(bottom)** panel, the FHR determined from the non-invasive fetal ECG (FHR_{NI}) with the proposed methods (corresponding to "This work" in **Tables 2**–**4**) is shown.

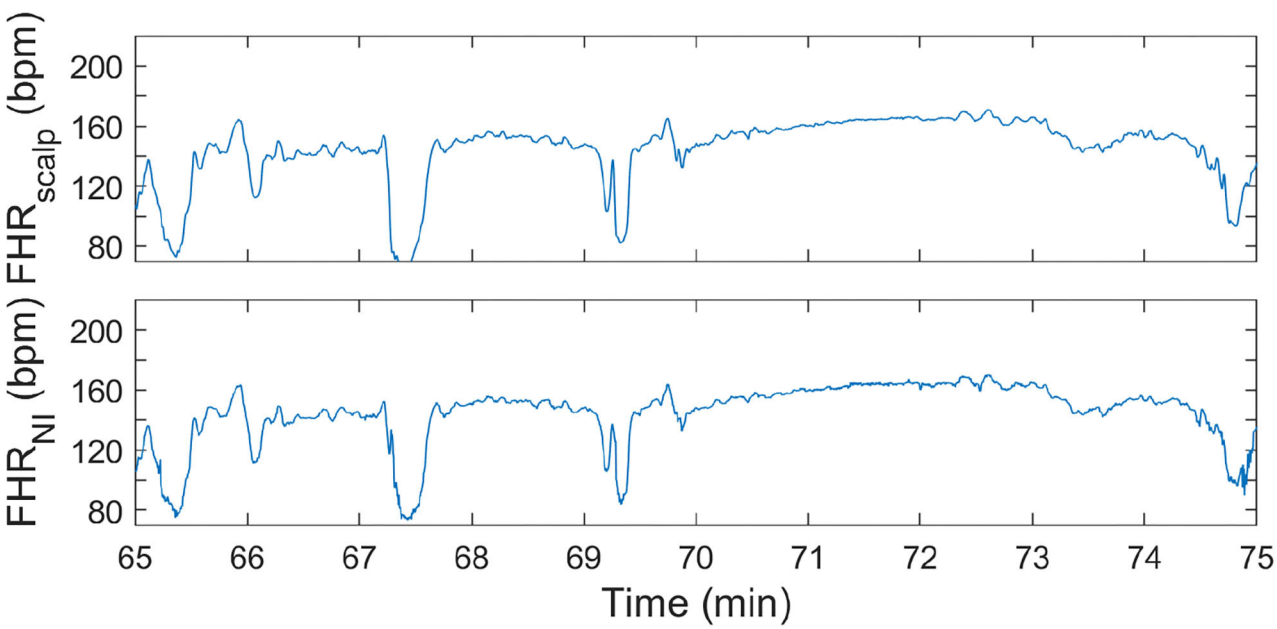

FIGURE 5 | FHR tracing during second stage of labor. In the **(top)** panel, the FHR from FSE is depicted. In the **(bottom)** panel, the FHR determined from the non-invasive fetal ECG (FHR_{NI}) with the proposed methods (corresponding to "This work" in **Tables 2**–**4**) is shown.

For the comparison between the presented method and other non-invasive FHR methods, the discussion below focuses on the results that are presented in the column "This work (validation)" in **Tables 2–4**. In this comparison, the presented method shows overall, and during the first stage of labor, significantly higher success rate and reliability, with an accuracy that is comparable to other electrophysiology-based methods and that is better than that of Doppler ultrasound. It should be noted here that the quantitative measure for accuracy is chosen according to the literature (6, 12) and defined such that the lower the value, the better the accuracy.

During the second stage of labor, success rate is significantly higher, reliability is more than 10% higher than that of Monica AN24, but the accuracy is only slightly better than that of Monica AN24 and of the previous version of the Nemo Healthcare product (i.e., NI-fECG in **Tables 2–4**). The main reason for this relatively smaller yield in accuracy is that the presented method has a success rate of close to 100%. During strong contractions with active pushing of the mother, the signal quality of the electrophysiological data is reduced significantly and the chance of providing inaccurate results is therefore higher. Moreover, during these episodes the FHR typically decelerates to <100 bpm. As can be seen in **Figure 5**, the proposed method also shows these decelerations but they are typically slightly less pronounced, yielding the same clinical picture, but at the same time producing difference between FHRs that are in the range of 10–15 bpm. In comparison, the Monica AN24 and the previous version of the Nemo Healthcare monitor show lower success rates which in practice means that during these decelerations, when the signal quality is lower, they do not show FHR, ensuring that inaccurate FHR during these episodes does not lead to erroneous interpretation but also that these inaccurate FHRs do not accumulate in a even further reduced accuracy.

TABLE 2 | Performance of various methods for FHR detection, as compared to the FHR from FSE as ground truth.

	Overall				
Metric	**This work (validation)**	**This work (train)**	**NI-fECG**	**Monica AN24**	**Ultrasound**
Success rate (%)	99.9 ± 0.2	99.6 ± 3.4	89.5 ± 10.8	83.4 ± 20.1	82.5 ± 21.1
CI (%)	99.9–100.0	99.5–99.6	87.9–91.1	78.8–87.9	77.8–87.3
Reliability (%)	95.7 ± 4.3	95.7 ± 6.2	86.8* ± 16.3	81.7 ± 20.5	73.0 ± 24.6
CI (%)	95.3–96.0	95.6–95.8	84.2–89.5*	77.1–86.4	67.4–78.5
Accuracy (bpm)	3.2 ± 1.4	3.0 ± 0.6	−1.5** ± 4.2	5.3 ± 2.4	10.9 ± 5.8
CI (bpm)	3.1–3.3	2.9–3.0	−3.4–0.5	4.7–5.8	9.6–12.2
LoA (bpm)	−9.2–15.5	−9.1–15.0	−29.2–26.3	Not provided	Not provided

*The results are aggregated over all patients and entire recordings. Confidence intervals for success rate and reliability are truncated at 100.0% and confidence intervals for accuracy are truncated at 0.0 (except for NI-fECG due to the different accuracy metric used). *In the paper by (12) on NI-fECG a margin of 10 bpm difference with FHR from FSE was used, instead of a margin of 10%, to assess reliability.*

***In the paper by (12) on NI-fECG the accuracy is determined as the average of the signed differences between the FHR from NI-fECG and the FHR from FSE instead of absolute differences. CI, 95% confidence interval; LoA, limits of agreement.*

TABLE 3 | Performance of various methods for FHR detection, as compared to the FHR from FSE as ground truth.

	Stage 1				
Metric	**This work (validation)**	**This work (train)**	**NI-fECG**	**Monica AN24**	**Ultrasound**
Success rate	99.9 ± 0.2	99.5 ± 3.7	91.3 ± 9.9	86.4 ± 21.1	82.6 ± 24.4
CI	99.5–100.0	92.1–100.0	89.8–92.8	81.6–91.2	77.0–88.2
Reliability	96.0 ± 3.4	95.7 ± 6.4	88.4* ± 14.6	84.9 ± 21.5	74.7 ± 28.2
CI	88.8–100.0	83.0–100.0	86.0–90.8*	80.0–89.8	68.2–81.2
Accuracy (bpm)	3.0 ± 1.8	2.9 ± 0.6	−1.4** ± 3.7	4.5 ± 2.4	7.9 ± 4.2
CI (bpm)	0.0–6.8	1.7–4.0	−3.2–0.4	3.9–5.0	7.4–10.0
LoA (bpm)	−8.3–14.3	−8.5–14.3	−27.2–24.4	−8.7–8.4	−28.4–22.7

*The results are aggregated over all patients and for first stage of labor of the recordings. Confidence intervals for success rate and reliability are truncated at 100.0% and confidence intervals for accuracy are truncated at 0.0 (except for NI-fECG due to the different accuracy metric used). *In the paper by (12) on NI-fECG a margin of 10 bpm difference with FHR from FSE was used, instead of a margin of 10%, to assess reliability. **In the paper by (12) on NI-fECG the accuracy is determined as the average of the signed differences between the FHR from NI-fECG and the FHR from FSE instead of absolute differences. CI, 95% confidence interval; LoA, limits of agreement.*

TABLE 4 | Performance of various methods for FHR detection, as compared to the FHR from FSE as ground truth.

	Stage 2				
Metric	**This work (validation)**	**This work (train)**	**NI-fECG**	**Monica AN24**	**Ultrasound**
Success rate	99.9 ± 0.1	99.8 ± 0.2	63.3 ± 21.7	75.2 ± 19.2	77.8 ± 21.1
CI	99.5–100.0	99.4–100.0	58.7–67.8	69.4–81.1	71.4–84.1
Reliability	85.9 ± 8.6	81.5 ± 11.4	68.5* ± 24.5	71.9 ± 20.4	61.7 ± 24.8
CI	58.5–100.0	58.1–100.0	62.9–74.1*	65.7–78.1	54.2–69.2
Accuracy (bpm)	6.6 ± 6.3	9.4 ± 2.7	−1.7** ± 8.2	7.9 ± 4.2	16.1 ± 7.6
CI (bpm)	0.0–26.6	4.0–14.9	−5.4–2.0	6.6–9.2	13.8–18.5
LoA (bpm)	−11.4–24.6	−19.1–37.9	−42.4–39.0	−12.3–12.4	−40.9–34.0

The results are aggregated over all patients and second stage of labor of the recordings. Confidence intervals for success rate and reliability are truncated at 100.0% and confidence intervals for accuracy are truncated at 0.0 (except for NI-fECG due to the different accuracy metric used).

**In the paper by (12) on NI-fECG a margin of 10 bpm difference with FHR from FSE was used, instead of a margin of 10%, to assess reliability.*

***In the paper by (12) on NI-fECG the accuracy is determined as the average of the signed differences between the FHR from NI-fECG and the FHR from FSE instead of absolute differences.*

CI, 95% confidence interval; LoA, limits of agreement.

4.1. Limitations

This study has four main limitations. First, in this study we have followed a common approach in evaluating the performance of AI methods by using a holdout dataset for validation. This holdout set can demonstrate the generalizability of the trained AI. In this study, however, the size of the validation set is relatively small. To still provide some insights in the potential of the proposed method on a larger dataset, we have included the results on the training data in **Tables 2–4**. While these results might be overestimating the performance of the proposed method due to overfitting, it should be noted here that the chance of overfitting is small. Not only are the results of the validation set similar to those of the training set, but also are the results of the AI not used directly to determine the FHR. More specifically, the results of the AI are used in the *prior* model of the hierarchical R-peak detection method and additional models—that e.g., consider the morphology of the signal at the expected position of the R-peak—are employed that likely prevent the detection of overfitted FHR values.

A second limitation of the study is that the comparison to other methods is indirect. The various methods have each been evaluated on their own data sets, with different numbers of patients, different patient characteristics, and different lengths of recordings. Therefore, no quantitative comparison can be made and no strong conclusions can be drawn about the performances of all methods. Yet, in our opinion the number of recordings in each study and the difference in performances is large enough to argue that the presented method outperforms the other methods in FHR detection during all stages of labor.

A third limitation is that the presented method, but also all reference methods, provide a FHR at 4 Hz sampling intervals. Fetal electrophysiological recordings potentially enable the study of beat-to-beat variability in the FHR which has been reported to yield better performance in detecting fetuses in distress when using linear features of fetal heart variability (3, 29). On the other hand, data resampled to 4 Hz has been reported to yield similar effects on features that reveal physiological changes during progression of labor and even better performance on detecting fetuses in distress when using entropy features of heart rate variability (29). Moreover, the communication protocols for most central monitoring systems require FHR values to be communicated at fixed frequency of 4 Hz. Because of the absence of a clearly best method to communicate FHR values (i.e., on a beat-to-beat basis or at a fixed frequency) and to adhere to the existing communication protocols and bring the methods presented in this paper already one step closer to implementation in clinical practice, we have chosen to equidistantly resample our data to 4 Hz.

The fourth limitation of the study is that all data processing presented in this study was done offline, on a desktop computer. While the non-AI parts of the method can be processed online (i.e., processing e.g., 1 s of data takes $<$1 s) on a normal desktop computer, the AI extension takes on average 2 s to process 1 s of data on a GPU (Titan V, NVIDIA, USA), when implemented in Tensorflow-Keras. Related to this limitation, unlike the reference methods shown in **Tables 2–4**, the presented method is not yet implemented in a medical device. Efforts to achieve this are currently ongoing.

4.2. Future Potential

Other than the implementation of the presented methods in a clinical device for reliable and unobtrusive FHR monitoring, the presented methods might have further potential to support obstetrical healthcare. Because the transabdominal recordings can, with relatively small additional effort, also provide the fetal ECG (30), further analysis of the ECG morphology, such as ST analysis might be possible. For ST analysis, accurate normalization of the fetal orientation would be crucial however. Fetuses in cephalic, transverse, or breech position would give different ECG morphology. We have shown in a previous study (31) that a different fetal orientation, or a different orientation of the electrical heart axis with respect to the abdominal electrodes, affects the degree of ST elevation and as such might affect ST alarms triggered on an obstetrical ward. Normalization for the fetal orientation would be possible by using ultrasound imaging (30) or different (i.e., relative) ST alarm mechanisms (32).

5. CONCLUSIONS

In this paper, a new method for FHR detection from non-invasive, transabdominal electrophysiological measurements was presented. The method is able to determine a reliable FHR in $>$95% of time during labor, making it substantially more reliable and accurate than Doppler ultrasound—the current clinical standard for non-invasive cardiotocography. During second stage of labor, the performance of the method decreases, but with a reliability higher than 80% it still outperforms Doppler ultrasound and other reference methods by a significant amount.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Data are available from the Data Governance Board of the Máxima Medical Center for researchers who can demonstrate they are qualified to use confidential data. Requests to access these datasets should be directed to Rik Vullings, r.vullings@tue.nl, Judith O. E. H. van Laar, judith.van.laar@mmc.nl.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Máxima Medical Center institutional review board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RV was responsible for data processing and analysis, statistical analysis, and drafting of the manuscript. JL was responsible for data collection and reviewing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The data collection of this research was funded by a Horizon 2020 grant and supported by Nemo Healthcare BV, the Netherlands (grant number 719500). Nemo Healthcare is a spin-off company from the Eindhoven University of Technology. Nemo Healthcare played no role in the analysis of the data, the drafting of the manuscript, or the decision to publish.

REFERENCES

1. World Health Organization. *Neonatal and Perinatal Mortality: Country, Regional and Global Estimates*. WHO Press (2006). Available online at: http://whqlibdoc.who.int/publications/2006/9241563206_eng.pdf
2. Vos A, Denktas S, Borsboom GJJM, Bonsel GJ, Steegers EAP. Differences in perinatal morbidity and mortality on the neighbourhood level in Dutch municipalities: a population based cohort study. *BMC Pregnancy Childbirth*. (2015) 15:201. doi: 10.1186/s12884-015-0628-7
3. Van Laar JOEH, Peters CHL, Vullings R, Houterman S, Bergmans JWM, Oei SG. Fetal autonomic response to severe acidaemia during labour. *BJOG*. (2010) 117:429–37. doi: 10.1111/j.1471-0528.2009.02456.x
4. Kawakita T, Reddy U, Landy H, Iqbal S, Huang CC, Grantz K. Neonatal complications associated with use of fetal scalp electrode: a retrospective study. *BJOG*. (2016) 123:1797–803. doi: 10.1111/1471-0528.13817
5. Euliano TY, Darmanjian S, Nguyen MT, Busowski JD, Euliano N, Gregg AR. Monitoring fetal heart rate during labor: a comparison of three methods. *J Pregn*. (2017) 2017:8529816. doi: 10.1155/2017/8529816
6. Cohen WR, Ommani S, Hassan S, Mirza FG, Solomon M, Brown R, et al. Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes: maternal surface electrode fetal monitoring. *Acta Obstetr Gynecol Scand*. (2012) 91:1306–13. doi: 10.1111/j.1600-0412.2012.01533.x
7. Hamelmann P, Vullings R, Schmitt L, Kolen AF, Mischi M, van Laar JOEH, et al. Improved ultrasound transducer positioning by fetal heart location estimation during Doppler based heart rate measurements. *Physiol Meas*. (2017) 38:1821–36. doi: 10.1088/1361-6579/aa8a1a
8. Reinhard J, Hayes-Gill BR, Schiermeier S, Hatzmann W, Herrmann E, Heinrich TM, et al. Intrapartum signal quality with external fetal heart rate monitoring: a two way trial of external Doppler CTG ultrasound and the abdominal fetal electrocardiogram. *Arch Gynecol Obstet*. (2012) 286:1103–7. doi: 10.1007/s00404-012-2413-4
9. Euliano TY, Nguyen MT, Marossero D, Edwards RK. Monitoring contractions in obese parturients: electrohysterography compared with traditional monitoring. *Obstetr Gynecol*. (2007) 109:1136–40. doi: 10.1097/01.AOG.0000258799.24496.93
10. Sameni R, Clifford GD. A review of fetal ECG signal processing; issues and promising directions. *Open Pacing Electrophysiol Ther J*. (2010) 3:4–20. doi: 10.2174/1876536X01003010004
11. Cohen WR, Hayes-Gill B. Influence of maternal body mass index on accuracy and reliability of external fetal monitoring techniques. *Acta Obstetr Gynecol Scand*. (2014) 93:590–5. doi: 10.1111/aogs.12387
12. Lempersz C, Noben L, Osta Gv, Wassen MLH, Meershoek BPJ, Bakker P, et al. Intrapartum non-invasive electrophysiological monitoring: a prospective observational study. *Acta Obstetr Gynecol Scand*. (2020) 99:1387–95. doi: 10.1111/aogs.13873
13. Kanjilal PP, Palit S, Saha G. Fetal ECG extraction from single-channel maternal ECG using singular value decomposition. *IEEE Trans Biomed Eng*. (1997) 44:51–9.
14. Sameni R. *Extraction of Fetal Cardiac Signals From an Array of Maternal Abdominal Recordings*. Institut National Polytechnique de Grenoble – INPG; Sharif University of Technology (2008).
15. Vullings R, Peters CHL, Sluijter RJ, Mischi M, Oei SG, Bergmans JWM. Dynamic segmentation and linear prediction for maternal ECG removal in antenatal abdominal recordings. *Physiol Meas*. (2009) 30:291–307. doi: 10.1088/0967-3334/30/3/005
16. Varanini M, Tartarisco G, Billeci L, Macerata A, Pioggia G, Balocchi R. An efficient unsupervised fetal QRS complex detection from abdominal maternal ECG. *Physiol Meas*. (2014) 35:1607–19. doi: 10.1088/0967-3334/35/8/1607
17. Andreotti F, Riedl M, Himmelsbach T, Wedekind D, Wessel N, Stepan H, et al. Robust fetal ECG extraction and detection from abdominal leads. *Physiol Meas*. (2014) 35:1551–67. doi: 10.1088/0967-3334/35/8/1551
18. 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:R1–35. doi: 10.1088/0967-3334/37/5/R1
19. Warmerdam GJJ, Vullings R, Schmitt L, Van Laar JOEH, Bergmans JWM. Hierarchical probabilistic framework for fetal R-peak detection, using ECG waveform and heart rate information. *IEEE Trans Signal Process*. (2018) 66:4388–97. doi: 10.1109/TSP.2018.2853144
20. 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) 29:829–41. doi: 10.1088/0967-3334/29/7/011
21. Euliano TY, Nguyen MT, Darmanjian S, McGorray SP, Euliano N, Onkala A, et al. Monitoring uterine activity during labor: a comparison of 3 methods. *Am J Obstetr Gynecol*. (2013) 208:66.e1–6. doi: 10.1016/j.ajog.2012.10.87
22. 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 Obstetr Gynecol Reprod Biol*. (2017) 215:197–205. doi: 10.1016/j.ejogrb.2017.05.027
23. Warmerdam GJJ, Vullings R, Schmitt L, Van Laar JOEH, Bergmans JWM. A fixed-lag Kalman smoother to filter power line interference in electrocardiogram recordings. *IEEE Trans Biomed Eng*. (2017) 64:1852–61. doi: 10.1109/TBME.2016.2626519
24. Rooijakkers MJ, Rabotti C, Oei SG, Mischi M. Low-complexity R-peak detection for ambulatory fetal monitoring. *Physiol Meas*. (2012) 33:1135–50. doi: 10.1088/0967-3334/33/7/1135
25. Fotiadou E, Xu M, van Erp B, van Sloun R, Vullings R. Deep convolutional long short-term memory network for fetal heart rate extraction. In: *IEEE Proceedings on Engineering in Medicine and Biology, Vol. 42*, Montréal, QC (2020).
26. Peters CHL, Vullings R, Rooijakkers MJ, Bergmans JWM, Oei SG, Wijn PFF. A continuous wavelet transform-based method for time-frequency analysis of artefact-corrected heart rate variability data. *Physiol Meas*. (2011) 32:1517–27. doi: 10.1088/0967-3334/32/10/001
27. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals. *Circulation*. (2000) 101:e215–20. doi: 10.1161/01.CIR.101.23.e215
28. Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat*. (1979) 7:1–26.
29. Goncalves H, Costa A, Ayres-de Campos D, Costa-Santos C, Rocha AP, Bernardes J. Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability. *Med Biol Eng Comput*. (2013) 51:665–76. doi: 10.1007/s11517-013-1036-7

30. Lempersz C, van Laar JO, Clur SAB, Verdurmen KM, Warmerdam GJ, van der Post J, et al. The standardized 12-lead fetal electrocardiogram of the healthy fetus in mid-pregnancy: a cross-sectional study. *PLoS ONE.* (2020) 15:e0232606. doi: 10.1371/journal.pone.0232606
31. Vullings R, Verdurmen KMJ, Hulsenboom ADJ, Scheffer S, de Lau H, Kwee A, 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:e0175823. doi: 10.1371/journal.pone.0175823
32. Hulsenboom ADJ, Verdurmen KMJ, Vullings R, van der Hout-van der Jagt MB, Kwee A, van Laar JOEH, et al. Relative versus absolute rises in T/QRS ratio by ST analysis of fetal electrocardiograms in labour: a case-control pilot study. *PLoS ONE.* (2019) 14:e0214357. doi: 10.1371/journal.pone.0214357

Conflict of Interest: RV is co-founder and share holder in Nemo Healthcare BV, the Netherlands.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Artificial Intelligence

ORIGINAL RESEARCH
published: 08 March 2021
doi: 10.3389/frai.2021.622616

A Machine Learning Approach to Monitor the Emergence of Late Intrauterine Growth Restriction

Nicolò Pini[1,2], Maristella Lucchini[2], Giuseppina Esposito[3], Salvatore Tagliaferri[3], Marta Campanile[3], Giovanni Magenes[4†] and Maria G. Signorini[1†]*

[1]Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Milan, Italy, [2]Department of Psychiatry, Columbia University Irving Medical Center, New York, NY, United States, [3]Department of Obstetrical Gynaecological and Urological Science, Federico II University, Napoli, Italy, [4]Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy

Late intrauterine growth restriction (IUGR) is a fetal pathological condition characterized by chronic hypoxia secondary to placental insufficiency, resulting in an abnormal rate of fetal growth. This pathology has been associated with increased fetal and neonatal morbidity and mortality. In standard clinical practice, late IUGR diagnosis can only be suspected in the third trimester and ultimately confirmed at birth. This study presents a radial basis function support vector machine (RBF-SVM) classification based on quantitative features extracted from fetal heart rate (FHR) signals acquired using routine cardiotocography (CTG) in a population of 160 healthy and 102 late IUGR fetuses. First, the individual performance of each time, frequency, and nonlinear feature was tested. To improve the unsatisfactory results of univariate analysis we firstly adopted a Recursive Feature Elimination approach to select the best subset of FHR-based parameters contributing to the discrimination of healthy vs. late IUGR fetuses. A fine tuning of the RBF-SVM model parameters resulted in a satisfactory classification performance in the training set (accuracy 0.93, sensitivity 0.93, specificity 0.84). Comparable results were obtained when applying the model on a totally independent testing set. This investigation supports the use of a multivariate approach for the *in utero* identification of late IUGR condition based on quantitative FHR features encompassing different domains. The proposed model allows describing the relationships among features beyond the traditional linear approaches, thus improving the classification performance. This framework has the potential to be proposed as a screening tool for the identification of late IUGR fetuses.

Keywords: late intrauterine growth restriction, machine learning, perinatal medicine, predictive monitoring, support vector machines

OPEN ACCESS

Edited by:
Martin Gerbert Frasch, University of Washington, United States

Reviewed by:
Ahsan H. Khandoker, Khalifa University, United Arab Emirates
Raghvendra Mall, Qatar Computing Research Institute, Qatar
Beth J. Allison, Hudson Institute of Medical Research, Australia

****Correspondence:***
Nicolò Pini
nicolo.pini@polimi.it

†These authors share last authorship

Specialty section:
This article was submitted to Medicine and Public Health, a section of the journal Frontiers in Artificial Intelligence

Received: *28 October 2020*
Accepted: *18 January 2021*
Published: *08 March 2021*

Citation:
Pini N, Lucchini M, Esposito G, Tagliaferri S, Campanile M, Magenes G and Signorini MG (2021) A Machine Learning Approach to Monitor the Emergence of Late Intrauterine Growth Restriction. Front. Artif. Intell. 4:622616. doi: 10.3389/frai.2021.622616

INTRODUCTION

Antenatal fetal heart rate (FHR) is a widely used tool to monitor fetal wellbeing (Chen et al., 2011). The assessment of fetal heart rate variability (HRV) has been reported to inform on the functional state of the autonomic nervous system (ANS), thus providing an indication on the fetal development throughout pregnancy. In the context of fetal pathological states, intrauterine growth restriction (IUGR) is one of the most relevant complications of pregnancy and it has been reported to alter HRV

(Huhn et al., 2011; Signorini et al., 2020b). IUGR is associated with a decreased rate of fetal growth, which is the result of an abnormal supply of maternal nutrients and placental transfer to the fetus. IUGR is a pathological fetal state characterized by an increased mortality and/or morbidity (Rosenberg, 2008; Smith et al., 2013; Sharma et al., 2016). The two phenotypes of IUGR (early and late) can be identified based on onset, evolution, Doppler parameters modifications, and postnatal outcome (Esposito et al., 2019).

In this paper we will focus on a population of late IUGR which is a condition with substantial increased prevalence if compared to early IUGR (Villar et al., 2014; Gordijn et al., 2016). The main cause for the insurgence of late IUGR is fetal hypoxemia/hypoxia secondary to placental insufficiency. Moreover, it is often associated with multiple placental anomalies that by contrast have less influence on placental resistance. Therefore, the umbilical Doppler indices are often unaffected, thus making the diagnosis of late IUGR more difficult, due to the large variability of fetal parameters on growth charts in the third trimester (Mureşan et al., 2016). Late IUGR is suspected when the fetal growth curve slows down or does not physiologically increase as a function of gestational age. Undetected IUGR in the third trimester of pregnancy represents the main cause of unexplained stillbirths in low-risk pregnancies, thus better antenatal diagnosis and treatment and timely delivery could diminish the risks significantly (Warland and Mitchell, 2014). In order to investigate identification of late IUGR through FHR analysis, we used the cardiotocography (CTG), which combines the measure of FHR through a Doppler ultrasound probe with the detection of uterine contractions using a pressure sensor. Although CTG analysis is still performed visually in the majority of Ob-Gyn clinical settings [following guidelines edited by national and international scientific societies, such as the International Federation of Gynecology and Obstetrics (FIGO) (FIGO, 1986)], a progressive transition to computerized approaches has been reported in recent years. Computerized systems are able to extract FHR parameters from multiple domains [time domain and frequency domain, complexity/nonlinear methodologies (Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology, 1996)] and represent the initial step toward multiparametric and multidimensional FHR analyses able to benefit from machine learning algorithms. As a matter of fact, the FHR regulation is the result of multiple and diverse neurological feedback loops, hormones, and various external factors, thus resulting into complex temporal dynamics, which are usually missed by the simple visual inspection. Additionally, previous studies have shown the strength of a multivariate framework in detecting fetal acidemia (Spilka et al., 2017) and a previous paper from our group used machine learning approaches to diagnose IUGR, but mainly focusing on the early phenotype (Signorini et al., 2020b).

In this study, classical FHR features were complemented with advanced nonlinear features and subsequently employed to train several machine learning algorithms for the detection of late IUGR in a database of 262 fetuses. Results highlighted the enhanced performance of nonlinear features over traditional parameters and the significant improvement in specificity and sensitivity of multiparametric machine learning approaches over univariate analysis. Furthermore, by utilizing an interpretable variant of support vector machines, we were able to identify the features that contributed the most to classification accuracy. This implementation provided meaningful and interpretable results with the potential of their translation into clinical practice.

MATERIAL: DATABASE AND PREPROCESSING

Dataset

Antepartum FHR recordings were collected at the Azienda Ospedaliera Universitaria—Federico II (Napoli, Italy). Data analyzed in the investigation were collected as part of the routine clinical examinations administered to pregnant women by the Italian healthcare system. Pregnant women signed informed consent for the utilization of their data for research purposes. The ethical committee and the IRB of Azienda Ospedaliera Universitaria—Federico II, Napoli, Italy, approved the enrollment of pregnant women as participants in the study and the utilization of the routine examination for research purposes. Traces were recorded in a controlled clinical environment, with participants lying supine on a bed undergoing a standard nonstress test protocol. Data were acquired using Philips cardiotocography (CTG) fetal monitor Avalon FM30 connected to a computer. The device employs an autocorrelation technique to compare the demodulated Doppler signal of a given heartbeat and the subsequent one. The resulting resolution for beat detection is below 2 ms. The derived CTG signal consists of a series of FHR values sampled at 2 Hz and expressed in beats per minute (bpm). Additionally, each FHR sample is accompanied by an indication of signal quality: optimal, acceptable, or insufficient based on the results of autocorrelation technique. We excluded pregnant women with maternal health conditions known to affect FHR regulation (gestational diabetes, preeclampsia, and hypertension), psychiatric medication use during pregnancy (SSRIs, antidepressants, classic antipsychotics, atypical antipsychotics, mood stabilizers, stimulants, antianxiety medications, or anticonvulsants), any recreational drug use during pregnancy, and congenital heart anomalies. Fetuses with congenital heart anomalies and genetic disorders were also excluded. The cohort analyzed in this work comprised 102 late intrauterine growth restriction (IUGR) fetuses and 160 healthy fetuses matched for GA at the first CTG examination. Fetuses in both groups underwent a routine ultrasound examination at approximately 34 weeks GA which did not exhibit any alteration in fetal growth or abnormalities in neither the middle cerebral artery nor the ductus venosus Doppler velocimetry. Once subsequently admitted for a CTG recording at 37.54 ± 0.77 (mean ± std) weeks, fetuses in the healthy group did not show any abnormality in the FHR trace, whereas IUGR fetuses (admitted at 36.94 ± 0.59 weeks) did present irregularities in their CTG recordings. A concurrent ultrasound examination showed alteration in both growth and Doppler profiles in this group of fetuses. The clinical definition of

TABLE 1 | Clinical data for the healthy and IUGR populations.

	Healthy n = 160	IUGR n = 102
GA at CTG (weeks)†	37.54 ± 0.77	36.94 ± 0.59
Maternal age (years)	32.23 ± 5.16	32.36 ± 5.82
Birthweight (g)†	3,311.62 ± 373.87	2,038.40 ± 348.15
Umbilical cord pH†	7.28 ± 0.08	7.32 ± 0.06
Fetal sex (male)	55.00%	46.08%
Apgar score 1 min >7†	91.88%	78.43%
Apgar score 5 min >7	100.00%	98.04%
Mode of delivery†	59.38% vaginal 40.62% caesarean	28.43% vaginal 71.57% caesarean

Expressed as mean ± standard deviation or number (%). †denotes a significant difference between healthy and IUGR fetuses, $p < 0.05$.

late IUGR adopted in this work reflects the guidelines reported in Gordijn et al. (2016). The adopted classification combines standalone information from multiple domains—gestational age, congenital abnormalities, absolute size measurements, and functional parameters—as well as their interactions (Gordijn et al., 2016). Each prenatal fetal condition was verified after delivery to confirm group membership previously suspected at the CTG timepoint. The length of the CTG recordings considered in this study is 40 min in both populations. This guaranteed that FHR data were acquired during both activity and quiet periods for any given fetus included in the study. Participants were not included in the analysis if their associated recordings were of insufficient duration and/or with <30 usable 1-min epochs and/or <10 3-min epochs (see next section for definition of epochs). Clinical data of the analyzed populations are reported in **Table 1**. Fetuses in the healthy group were characterized by higher birthweight, Apgar score 1 min, and rate of spontaneous vaginal delivery compared to the considered IUGR population. Similar results have been reported in other populations of early (Stampalija et al., 2015) and late (Esposito et al., 2019) IUGR fetuses.

Fetal Heart Rate Time Series and Preprocessing

The equipment used to record the data under investigation provides each FHR sample with an indication of signal quality: optimal (green), acceptable (yellow), or insufficient/absent (red) based on the results of the embedded autocorrelation procedure utilized to extract the signal itself (Lawson et al., 1983; Signorini et al., 2003). The first preprocessing step toward the computerized analysis of the acquired traces consisted in subdividing each FHR recording in shorter segments of length 120 points (60 s) or 360 points (180 s). The choice of 1-min or 3-min subintervals is related to the different time scales on which CTG-derived features are computed, whose procedure will be described in the following sections. Subsequently, segments including more than five consecutive red-quality points or more than 5% of red-quality samples (6 FHR values out of 120 points per subinterval or 18 FHR values out of 360 points per subinterval) were discarded in further analysis. Lastly, isolated insufficient-quality points were substituted, through a moving average procedure, with the average of the nearest five FHR points. For an in-depth description of the preprocessing steps adopted in this investigation, see previous publications by our group (Arduini et al., 1993; Signorini et al., 2003; Magenes et al., 2007).

METHODS: FEATURES AND RADIAL BASIS FUNCTION SUPPORT VECTOR MACHINES

Features

The present contribution focuses on building a machine learning-based screener of late IUGR pathology fed with a set of FHR features rather than a single feature design-oriented approach. Thus, the employed features were selected on the basis of the a priori knowledge on various quantifiers of fetal ANS dynamics in different domains, complemented by fetal and maternal information. **Figure 1** reports a schematic workflow for the framework implemented in this work.

Morphological and Time Domains

Morphological and time domain features represent the computerized and automated extraction of FIGO guidelines from FHR recordings, in terms of baseline evolution, accelerations/decelerations, and variability. Starting from the identification of FHR baseline [by means of Mantel's approach (Mantel et al., 1990)], it is possible to derive the automatic counts of large accelerations (more than 15 beats per minute over the baseline lasting 15 s or more) (*#acc_large*), small accelerations (fewer than 15 beats per minute) (*#acc_small*), decelerations (*#dec*), and contractions (*#contr*) (Rabinowitz et al., 1983; FIGO, 1986). FHR variability features are derived from FHR signal excluding events of accelerations and decelerations. Specifically, the overall variability is quantified by the mean and standard deviation of entire FHR signal (*FHR_mean* and *FHR_std*). Short Term Variability (*STV*), Interval Index (*II*), and *Delta* provide estimates of short term FHR variability considering 1-min FHR intervals. Long Term Irregularity (*LTI*) quantifies variability on a longer time scale (3-min FHR intervals). A more comprehensive description and characterization of the employed FHR variability features can be found in previous publications by our group (Arduini et al., 1993; Signorini et al., 2003; Magenes et al., 2007).

Frequency Domain

Power Spectral Density (PSD) is a largely exploited method for HRV frequency analysis. It decomposes the signal power in oscillatory components which are an indirect measure of ANS modulation over the cardiac system. PSD is computed employing autoregressive (AR) modeling, specifically by Levinson-Durbin algorithm. Based on previous findings, three specific frequency bands of interest can be identified, namely, low frequency (LF) (0.03–0.15) Hz; movement frequency MF (0.15–0.5) Hz; high frequency HF (0.5–1 Hz), which quantifies the different branches of ANS modulation (Signorini et al., 2003; Faes et al., 2015; Spilka et al., 2017). The FHR signal undergoes an automatic decomposition into a sum of sinusoidal contributions identified by their central frequencies and the associated

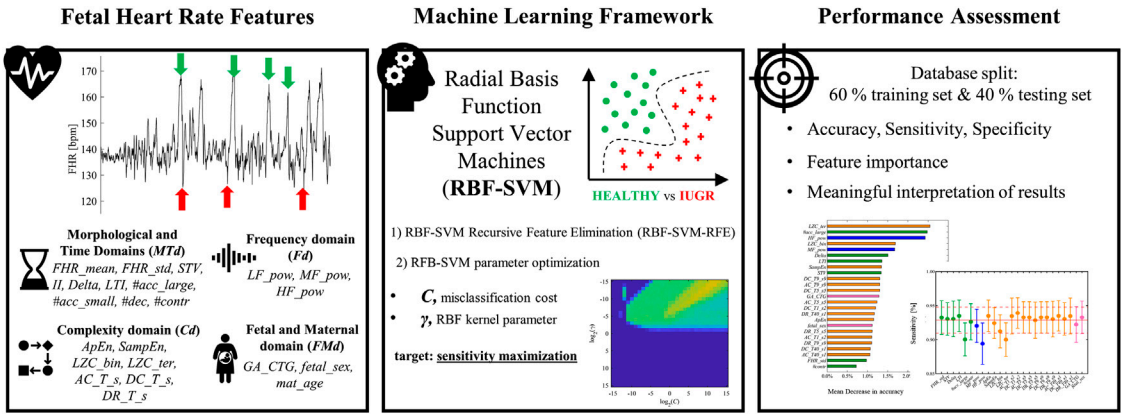

FIGURE 1 | A schematic depiction of the machine learning framework to monitor the emergence of late intrauterine growth restriction. The starting point was the extraction of fetal heart rate physiology features. Morphological and Time Domains *(MTd)* – mean, standard deviation of entire FHR signal *(FHR_mean, FHR_std)*; Short Term Variability *(STV)*; Interval Index *(II)*; Delta; Long Term Irregularity *(LTI)*; large accelerations (more than 15 beats per minute over the baseline lasting 15 s or more, indicated by green arrows) *(#acc_large)*; small accelerations (fewer than 15 beats per minute) *(#acc_small)*; decelerations (*#dec,* indicated by red arrows); contractions *(#contr)*. Frequency domain *(Fd)* – power in the low frequency (LF) band *(LF_pow)*; power in the movement frequency (MF) band *(MF_pow)*; power in the high frequency (HF) band *(HF_pow)*. Complexity domain *(Cd)*—Approximate Entropy *(ApEn)*; Sample Entropy *(SampEn)*; binary, ternary Lempel Ziv Complexity *(LZC_bin, LZC_ter)*; Average Acceleration, Deceleration Capacities *(AC_T_s, DC_T_s)*; Deceleration Reserve *(DR_T_s)*. Fetal and Maternal domain *(FMd)* – GA at the recording *(GA_CTG)*; fetal sex *(fetal_sex)*; maternal age *(mat_age)*. Radial basis function support vector machine (RBF-SVM) model was developed to discriminate between healthy and late IUGR fetuses. The ensemble of previously described features was reduced by means of Recursive Feature Elimination technique (RBF-SVM-RFE). The optimization of RBF-SVM parameters C and γ aimed at maximizing sensitivity is shown in the bottom graph of the Machine Learning Framework panel. Performance was assessed by splitting the database of 160 healthy and 102 late IUGR fetuses in a training (60%) and independent testing set (40%). Several figures of merit were computed. Additionally, the proposed implementation of RBF-SVM allowed deriving interpretable feature importance ranking and standalone feature contribution to accuracy, sensitivity, and specificity (left and right (only sensitivity shown) graphs in the Performance Assessment panel).

amount of power, thus obtaining the power in the LF band (*LF_pow*), MF band (*MF_pow*), and HF band (*HF_pow*) for each 3-min FHR segment.

Complexity Domain

The application of nonlinear methodologies to investigate FHR variability has demonstrated its usefulness in predicting fetal wellbeing in several investigations (Signorini et al., 2003; Spilka et al., 2012; Gonçalves et al., 2018). In the context of early IUGR detection, Lempel Ziv Complexity (LZC) (Lempel and Ziv, 1976) has been previously reported to have enhanced discriminative power in both univariate (Ferrario et al., 2007) and multivariate (Signorini et al., 2020b) approaches, considering binary (*LZC_bin*) and ternary (*LZC_ter*) alphabets. Additional measures of complexity such as Approximate Entropy (*ApEn*) (Pincus and Viscarello, 1992) and Sample Entropy (*SampEn*) (Lake et al., 2002) have been employed for the quantification of ANS profiles in the perinatal period. The above described features are computed for each 3-min FHR segment. The last nonlinear technique is Phase Rectified Signal Averaging (PRSA) (Bauer et al., 2006). Average Acceleration and Deceleration Capacities (*AC* and *DC*) are among the various parameters which can be derived from the PRSA curve (Fanelli et al., 2013). More recently, Deceleration Reserve (*DR*) (Rivolta et al., 2019) was defined as the simple summation of *AC* and *DC* and it has been shown to achieve enhanced performance in detecting fetal hypoxia compared to *AC* and *DC* standalone parameters in the context of intrapartum FHR recordings. Regarding the specific implementation of these methodologies in this work, for the computation of *LZC_bin* and *LZC_ter* the factor value (p) was set to zero, whereas for entropy computation the length of the pattern (m) and tolerance (r) were set equal 1 and 0.1, respectively, accordingly with the prior knowledge on their application for fetal investigations (Faes et al., 2015; Gonçalves et al., 2018; Signorini et al., 2020b). On the other hand, a technical aspect that complicates the physiological understanding of PRSA-derived features is their dependence on three parameters, namely, T, s, and L. Different combinations of the former parameters allow gaining insight about the ANS branches separately. In this work, *AC*, *DC*, and *DR* were computed considering $T = 1$ and $s = 2$, $T = 5$ and $s = 5$, $T = 9$ and $s = 9$, $T = 40$ and $s = 1$, and L was constant and equal 100.

Fetal and Maternal Domain

The evolution of fetal ANS regulation throughout pregnancy has been extensively investigated in regard to GA, sex, and various aspects (Giuliano et al., 2017; Gonçalves et al., 2017, 2018). This evidence is consistently reported among *MTd*, *Fd*, and *Cd* features. To address this issue, GA at the recording (*GA_CTG*), fetal sex (*fetal_sex*), and maternal age (*mat_age*) are included in the machine learning analyses.

Feature Preprocessing

The time series of each parameter was averaged throughout the recording to derive a single set of features for each subject. At this step, features were preprocessed for outliers [Winsorization in the interval ($Q_1 - 3IQR$, $Q_3 + 3IQR$), where Q_i is defined as the ith quartile and IQR = $Q_3 - Q_1$]. Lastly, features were standardized

across the entire population to obtain zero mean and unitary variance distributions.

Radial Basis Function Support-Vector Machines

From Linear to Radial Basis Function Support Vector Machines

Support Vector Machines (SVM) are a class of machine learning algorithms highly exploited for the purposes of data classification and regression. As a general consideration, SVM aim to derive a model learned on the training data, which is able to predict the target values contained in the test data (Cortes and Vapnik, 1995; Guyon et al., 2002; Hsu, Chang and Lin, 2003). Given a training set of labeled instance pairs $(\boldsymbol{x}_i, y_i)$, $i = 1,\ldots,l$, where $\boldsymbol{x}_i \in \mathbb{R}^n$ and $y \in \{1,-1\}^l$, l is equal to the number of observed pairs, n is the dimensionality of the feature space, and y corresponds to healthy/late IUGR binary classification assigned to each participant. SVM searches for the optimal hyperplane $\boldsymbol{w}^T\phi\ (\boldsymbol{x}_i) + b$, which maximizes the separation margin between the two classes by solving an optimization problem. In the context of classical SVM, such function is linear; thus the corresponding term reads $\boldsymbol{w}^T\boldsymbol{x}_i + b$, which translates into a linear separating hyperplane. $C > 0$ is the so-called penalty parameter of the error term. C controls the tradeoff between misclassification and data sparsity. Large values of C constrain the optimization procedure to derived smaller-margin hyperplane if such boundary contributes to the training points classified correctly. Conversely, a smaller value of C causes the optimizer to search for larger margins, even if the derived hyperplane misclassifies more observations. Classical SVM promote data sparsity given only few subjects contribute to the margin determination at the expenses of involving all the features, thus being affected by the curse of dimensionality (Hsu et al., 2003). To address the described issues, we propose to employ a more efficient kernel function: Radial Basis Function SVM (RBF-SVM) (Hsu et al., 2003), and a novel feature elimination algorithm, namely, RBF-SVM Recursive Feature Elimination (RBF-SVM-RFE) (Liu et al., 2011). The main shortcoming of classical SVM is the constraint of describing the relationship between the class labels and the features as linear. On the opposite, the kernel of RFB-SVM maps observations into a higher dimensional space, thus allowing for a nonlinear relationship between observations and attributes. In this scenario, the function ϕ can be expressed according to (**Eq. 1**):

$$K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = e^{-\gamma\boldsymbol{x}_i - \boldsymbol{x}_j^2}, \qquad \gamma > 0\,, \tag{1}$$

where K is called kernel function, and the parameter γ defines the radius of influence of a given training example. Specifically, low values of γ code for far influence and a very broad decision region, whereas high values of γ result in the opposite. Furthermore, it can be shown that RFB kernel is equivalent to the linear one for some combinations of (C, γ) (Lin and Lin, 2003). RBF-SVM are suitable to be employed in the presented study given the well-documented nonlinear relationship between several features and the target binary outcomes: healthy or IUGR fetuses (Signorini et al., 2003; Spilka et al., 2017; Gonçalves et al., 2018).

Radial Basis Function Support Vector Machine Recursive Feature Elimination

Linear SVM Recursive Feature Elimination (SVM-RFE) is a largely exploited category among the wrapper models (Kohavi and John, 1997) which performs feature selection (Guyon et al., 2002). Wrapper methodologies are computationally demanding but they exhibit enhanced performance compared to filter approaches (Sun, 2007). If SVM-RFE allows deriving an interpretable feature ranking, the same is not valid when considering nonlinear SVM (as for RBF-SVM). This relates to the fact the mapping function ϕ is unknown; thus the vector $\boldsymbol{w}$ cannot be explicitly computed and consequently cannot be used to rank features as for SVM-RFE. In this work, we employed a recent extension of SVM-RFE which performs feature elimination in the context of nonlinear SVM, namely, RBF-SVM Recursive Feature Elimination (RBF-SVM-RFE) (Liu et al., 2011). In a nutshell, RBF-SVM-RFE expands RBF kernel into its Maclaurin series. The weight vector $\boldsymbol{w}$ is derived from the series by computing the contribution made to the classification hyperplane of each feature. The algorithm allows ranking features by their relative importance starting by including all features and progressively eliminating each of them until all attributes are ranked. Moreover, RBF-SVM-RFE allows deriving the most informative subset of feature among all possible permutations of the original set. A comprehensive and rigorous description of the algorithm can be found in Liu et al. (2011).

Performance Assessment

Performance is quantified in terms of the area under receiver-operation-characteristic (ROC) curve (AUC), sensitivity (SE), and specificity (SP). In the context of supervised machine learning approaches as for RBF-SVM, it is usually required to perform the following: 1) make use of cross-validation (CV) to identify the best pair of parameters C and γ; 2) train the whole training set using the previously identified C^{opt} and γ^{opt} and evaluate the performance; 3) test the validity, replicability, and stability of the model on a new set of observations which have never been used in the training phase. The prediction accuracy obtained from the unknown observations is thought to reflect in a more precise way the classification performance of the trained algorithm. In the context of this work, the training set was obtained by including 60% of the original dataset and utilized to perform task 1) and task 2), whereas the remaining 40% was used to derive the independent test set and employed in task 3). The operation of testing the model performance uniquely on validation dataset does not guarantee unbiased results as the model is fitted on the training dataset while tuning model hyperparameters. On the contrary, the utilization of an independent test has been shown to provide an unbiased evaluation of the final model. The ratio between healthy and IUGR (~1.5:1) was maintained constant in both sets. C^{opt} and γ^{opt} were derived by performing a grid search on C and γ using cross-validation. Specifically, several pairs of (C, γ) were tested and the one achieving the best cross-validation figure of merit was chosen.

TABLE 2 | Univariate performance.

	AUC	Sensitivity	Specificity
LZC_bin (bits)	0.78	0.78	0.68
LZC_ter (bits)	0.78	0.88	0.57
#acc_large	0.72	0.84	0.48
AC_T9_s9 (bpm)	0.68	0.76	0.50
AC_T5_s5 (bpm)	0.67	0.66	0.58
FHR_std (bpm)	0.66	0.72	0.55
#Contr	0.65	0.66	0.58
LTI (ms)	0.63	0.66	0.58
Delta (ms)	0.62	0.83	0.40
STV (ms)	0.61	0.98	0.21

Feature cut-offs associated with the optimal values of sensitivity and specificity are derived based on Youden's index maximization.

Exponentially growing sequences of C and γ were employed in a the grid search framework (Hsu et al., 2003; Spilka et al., 2017) along with a 10-Fold CV repeated 5 times on the training set to identify C^{opt} and γ^{opt}. In this work, SE was identified as the figure of merit to be maximized to the aim of deriving a screening tool suitable to be employed in clinical practice for the identification of late IUGR. Despite its straightforward implementation compared to more advanced methodologies, a grid search approach has the advance of avoiding approximations by performing an exhaustive parameter search. Additionally, it can be easily parallelized since each (C, γ) pair is independent. On the opposite, iterative processes can hardly be run simultaneously (Hsu et al., 2003).

RESULTS

Univariate Analysis

For benchmark, the performance in discriminating healthy vs. IUGR fetuses for each of the previously described features was computed. Specifically, a set of logistic regression models were trained including each attribute individually. The optimal cut-off (c) for a given feature was derived by the maximization of Youden Index (Fluss et al., 2005) defined as $J = max_c\,(SE\,(c) + SP\,(c) - 1)$. J allows computing the optimal c and consequently the corresponding SE (c), SP (c), and AUC (c) values. **Table 2** reports the ten best performing features ranked by their AUCs. *Cd* features yield the best univariate classification results, followed by *MTd* ones. Notably, neither *Fd* nor *FMd* attributes have a significant individual power. The selected features clearly point to the importance of more sophisticated analyses of FHR, rather than the traditional time and frequency approaches. Despite satisfactory values of AUC, the corresponding SE and SP suggest the need for a multivariate framework in order to improve and balance the overall performance. Prior to multivariate classification, correlation among all pairs of features was performed: 1) short and longer term *MTd* features were moderately correlated; 2) short term variability measured in the different domains: *ApEn*, *SampEn*, *HF_pow*, *LZC_bin*, and *LZC_ter* was highly correlated as expected given their definitions; 3) *ApEn*, *SampEn*, *LZC* parameters did not exhibit any relationship with PRSA-derived features; 4) *ACs* and *DCs* at different scales exhibited marked negative correlations; 5) *DRs* were weakly positive correlated with the corresponding *DCs* but not with *ACs*.

Multivariate Analysis

The performance of several machine learning classifiers was tested against the proposed RBF-SVM methodology. An exhaustive description of the employed techniques is reported in a previous publication by our group (Signorini et al., 2020b). Following the procedure illustrated by Zhang et al. (Zhang et al., 2002), the ROC curves associated with these methodologies were independently compared to the results of RBF-SVM model in a paired design. Results showed that all the tested techniques were statistically inferior to the RBF-SVM model. Nonetheless, features were ranked similarly among the tested machine learning classifiers, supporting the robustness of the proposed physiology based heart rate indices.

Feature Selection

The original set of features comprised $n = 32$ attributes, of which $n = 10$ from *MTd*, $n = 3$ from *Fd*, $n = 16$ from *Cd*, and $n = 3$ from *FMd*. The first step prior to multivariate analysis was to reduce the feature space according to RBF-SVM-RFE, as described in the Methods section. The minimum and maximum allowed numbers of features for each subset were $n = 1$ and $n = 32$, respectively. Among the tested subsets, the selected one consisted of 25 retained features and seven eliminated. Specifically, the features with the least squared weights were *FHR_mean*, *II*, *#acc_small*, *#dec*, *LF_pow*, *DR_T1_s2*, and *mat_age*. This result was in accordance with the findings for the univariate analysis. Consistently, the dropped attributes exhibited poor discriminative performance as standalone parameters. Additionally, the results of the correlation analysis for *FHR_mean*, *II*, and *mat_age* highlighted their independence of any other variable included in this analysis. *LF_pow* and

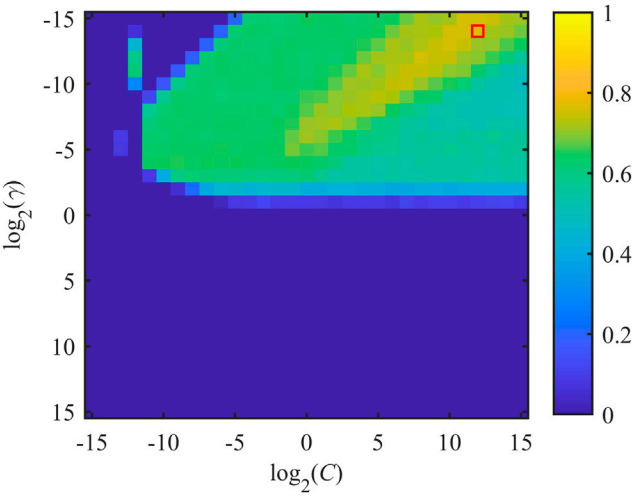

FIGURE 2 | Distribution of J as a function of misclassification cost C and SVM-RBF kernel parameter γ. x- and y-axes are expressed in logarithmic units for better interpretability of the adopted grid search. J^{opt} is achieved by considering the pair (C^{opt}, γ^{opt}), which is indicated by the red box.

DR_T1_s2 were highly correlated with frequency and PRSA-extracted indices; thus it is likely that their contribution in classification resulted as limited. Lastly, *#acc_small* and *#dec* did not exhibit substantial variations in the two groups.

Radial Basis Function Support Vector Machine Parameter Optimization

Various pairs of (C, γ) were tested to identify the combination (C^{opt}, γ^{opt}) which corresponded to the maximization of the figure of merit J. Each value of J was obtained by training the model on the whole set of selected features employing a 10-Fold CV scheme, repeated 10 times. Exponentially growing sequences of $C = \gamma = 2^{-15}, 2^{-14}, \ldots, 2^{14}, 2^{15}$ were adopted as practical implementation of the RBF-SVM grid search previously described. The distribution of J as a function of C and γ is shown in **Figure 2**. About half of the tested pairs (corresponding to positive exponents of γ) resulted in an unsatisfactory performance ($J = 0$), which is mapped in the lower half of the plane displayed in **Figure 2**. The remaining portion of the investigated two-dimensional space is associated with more satisfactory values of SE and SP. Specifically, the optimal combination was achieved by setting $C^{opt} = 2^{12}$ and $\gamma^{opt} = 2^{-14}$. In this case, J^{opt} was equal to 0.7682 and the corresponding SE and SP were equal to 0.9287 and 0.8395, respectively. Noticeably, SE associated with the reported J^{opt} is the highest achieved for the presented parameter searching. On the other hand, the best SP was equal to 0.8881 but the corresponding SE was 0.7467 ($C^{opt} = 2^5$ and $\gamma^{opt} = 2^{-15}$), thus being unsatisfactory from the perspective of building a screening tool.

Performance Assessment on Training and Testing Sets

The pair $C^{opt} = 2^{12}$ and $\gamma^{opt} = 2^{-14}$ was employed as optimal set of parameters for the final adopted model. This was learned on the training set by a 10-fold CV scheme repeated 10 times, including the restricted set of selected features. The resulting AUC was 0.9277 (0.9109, 0.9445), corresponding to SE equal 0.9287 (0.9095, 0.9479) and SP equal 0.8395 (0.8024, 0.8766). Results are reported as mean and 95% confidence interval. A main drawback of the proposed pipeline is the opportunity for overfitting the model on the training data. The practice of testing the derived model on a validation set aims at evaluating its robustness and insensitivity to overfitting. As previously described, the validation set encompasses 40% of the original dataset with the requirement of a similar ratio of healthy vs. IUGR cases. The model tested on the validation set achieved a close agreement with the one obtained on the training data. In detail, classification accuracy was 0.8462 (0.7622, 0.9094) and the associated values of SE and SP were 0.8438 and 0.8500, respectively. The resulting performance did not exhibit a drastic decrease of AUC, SE, or SP, strengthening the validity of the proposed model as a screening tool. This assumption was highlighted by additional figures of merit such Positive Predicted Value (PP V), 0.9000, and Negative Predictive Value (NPV), 0.7727.

Feature Importance

The main advantage in employing interpretable fetal heart rate features becomes evident for the purpose of providing meaningful machine learning findings. Specifically, the combination of heart rate attributes and RFB-SVM-RFE allows investigating the relative influence of each attribute toward classification. The results of described approach are displayed in **Figures 3**, **4**. The operation of ranking features according to the weight vector $\boldsymbol{w}$ was also found to reflect the mean decrease in accuracy of classification when a given feature was removed from the original set employed in the training phase, as shown in **Figure 3**. The features that, when removed, generated the biggest decrease in accuracy were found to belong to different domain, namely, *LZC_ter* for *Cd*, *#acc_large* for *MTd*, and *HF_pow* for *Fd*. Additionally, the reported mean decrease in discriminative power appeared limited if compared to the reference accuracy achieved in the training set. At the same time, the associated SE and SP highlighted a more pronounced decrease in performance as displayed in **Figure 4**. The feature specific decreases in SE shown in the left-hand panel in **Figure 4** were highly correlated with the results reported in **Figure 3**. In fact, *LZC_ter*, *#acc_large*, and *HF_pow* accounted for the greatest decrease in SE, whereas SE stayed stable once the remaining features were removed from the model. SP exhibited a similar behavior as reported in the right-hand panel in **Figure 4**. The described evidence suggests that the employed features contribute similarly to SE and SP. Based on these results, we can conclude that both SE and SP appear as robust figure of merit in the context of the proposed model.

DISCUSSION

In this investigation, we provided evidence for the successful application of a machine learning framework for the identification of late IUGR condition based on a single routine CTG examination. Starting from the unsatisfactory results of traditional univariate analysis (as reported in **Table 2**) we proposed an interpretable RBF-SVM model to be employed as screening tool in a clinical setting. The potential of early identification of late IUGR represents a noticeable step toward a better clinical management aimed at improving fetal outcome (Rosenberg, 2008). Discussing the model performance, the achieved values of AUC, and the associated SP and SE demonstrated the consistent ability of the proposed methodology to discriminate healthy vs. late IUGR fetuses in the training and in the validation set. This result is a consequence of the accurate tuning of model parameters (C, γ) designed to prevent overfitting. The proposed grid search for the optimal pair C, γ aimed at balancing the tradeoff between the values of model variance and bias. As a general consideration, high values of the misclassification cost (C) contribute to hard margin, thus forcing the model to a stricter interpretation of training data, potentially resulting in overfitting the training data. On the opposite, small values of gamma (γ) lead to low bias and high variance models. In this work, the selected pair consistently points to a high variance

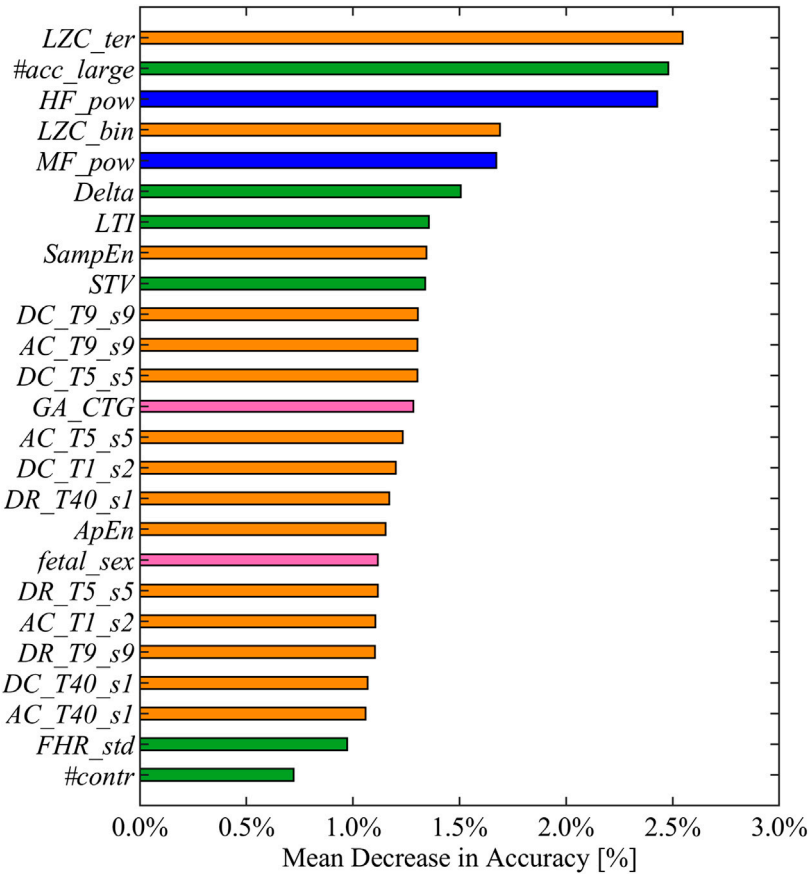

FIGURE 3 | Feature ranked by mean decrease in accuracy in descending order from top to bottom. Mean decreases in accuracy computed as the difference between the optimal accuracy (obtained by including the entire set of selected features in the training set) and the models learned excluding each feature alternatively. The displayed colors code for the different feature domains: green for *MTd*, blue for *Fd*, orange for *Cd*, and pink for *FMd*.

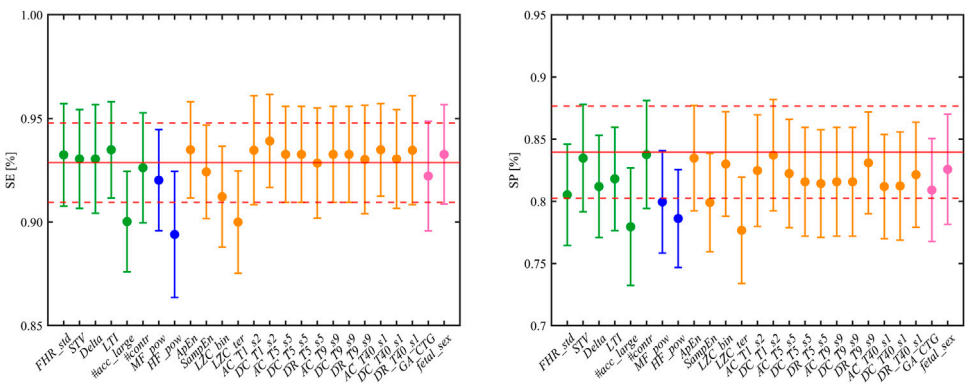

FIGURE 4 | Left and right panels show the resulting SE and SP (mean and CI) when each feature is alternatively excluded from the trained model. The reference values of SE and SP are reported in red, solid and dashed lines correspond to mean and CI, respectively. The displayed colors code for the different feature domains: green for *MTd*, blue for *Fd*, orange for *Cd*, and pink for *FMd*.

and low bias model. This translates in a separation hyperplane characterized by shaper boundaries and a strong penalization of misclassification error, suitable for the screening tool-oriented applications. Additionally, the absence of a potential bias toward overfitting is supported by the presented results on the validation set.

A crucial aspect for the clinical application of the presented model was the possibility to interpret the data-driven results assessing the features importance. The most noticeable advantage of RFB-SVM-RFE is its peculiar insight on the individual feature contribution to classification accuracy, SP, and SE. The feature ranking reported in **Figure 3** highlighted that the combination of features from different domains is effectively enhancing the model discriminative performance compared to traditional univariate analysis. Specifically, the top three features encompassed *Cd*, *Fd*, and *MTd* domains, respectively. This finding supports the notion that IUGR condition is effectively impairing the fetal ANS under different aspects; thus a comprehensive set of attributes are required for an accurate determination of such pathological condition. Moreover, the presented methodology allowed evaluating the contribution of each feature in terms of SE and SP as reported in **Figure 4**. If the SE contributes appeared moderately distributed among the features included in the model, this was not verified for SP. This figure of merit achieved the optimal performance only when all the selected attributes were included in the SVM framework. It is possible to speculate that the described behavior is a consequence of the grid search design. Specifically, the requirement of SE maximization allowed achieving an adequate SP at the expense of its robustness.

An additional advantage of RFB-SVM is the opportunity to define the relationship between features as nonlinear; thus, it allows overcoming the limitation of linearity imposed by traditional SVM approaches. At the same time, despite increasing the overall complexity of classification if compared to more traditional SVM implementations, the radial kernel tuning is on average of reduced complexity with respect to polynomial kernel given the fewer hyperparameters to be optimized. Lastly, RBF kernel is mathematically more stable in contrast to polynomial kernel which tends to converge to either infinity or zero for larger degrees (Hsu et al., 2003).

To our knowledge, this work is the first attempt toward a CTG and quantitative feature-based discrimination of late IUGR condition. Previous research mainly focused on the investigation of animal models (Poudel et al., 2015) and analyses of metabolic (Sanz-Cortés et al., 2013) and Doppler profiles (Parra-Saavedra et al., 2013) of chronic hypoxia in the fetal period. Nevertheless, the underpinning and widely reported consequence of long-lasting oxygen deprivation is responsible for a delay in the maturation of the branches of ANS and their subsequent integration with the central nervous system (CNS). The impairment in ANS maturation was consistently found in this investigation by various quantitative CTG-derived parameters which have been extensively associated with the fetal ANS modulation throughout pregnancy as standalone features (Signorini et al., 2003; Gonçalves et al., 2018). In comparison with previous machine learning-derived and univariate results by our group in different populations of early IUGR (Ferrario et al., 2007; Fanelli et al., 2013; Signorini et al., 2020b), it is possible to observe a consistent discriminative power of features *LZC*, *HF_pow*, and *LTI*. Specifically, the average value of each feature was greater in the control group vs. late IUGR fetuses. On the other hand, we also report an enhanced classification contribution of *SampEn*, which outperformed *ApEn*. Lastly, in the described late IUGR population, short scale ($T = s = 5$ and $T = s = 9$) PRSA-extracted features were characterized by a greater discriminative power compared to global ones ($T = 40$ and $s = 1$). The reported findings are in accordance with the univariate results and support the hypothesis of an impaired fetal beat-to-beat responsiveness regulation in the context of nutrient restriction and chronic hypoxemia (Fanelli et al., 2013; Rivolta et al., 2019). Lastly, toward enhancing the general applicability as well as interpretability of the proposed RBF-SVM model, we tested its performance once excluding the information of fetal sex. The knowledge about the sex of the fetus is banned in several countries across the globe; thus a fetal sex-independent model is expected to achieve wider applicability. Additionally, the influence of sex of the fetus over several of the physiology based heart rate features is object of open debate in the scientific community. Results showed nonstatistically different performances of fetal sex-removed RBF-SVM compared to the reference. Specifically, classification accuracy, sensitivity, and specificity were equal to 0.9208 (0.9012, 0.9413), 0.9247 (0.9018, 0.9493), and 0.7905 (0.7492, 0.8322); 0.8077 (0.7187, 0.8784), 0.8125, and 0.8000 in the training/testing and validation sets, respectively.

CONCLUSION

This contribution aims at promoting the application of machine learning methodologies in the context of fetal and perinatal medicine, following the growing trend of the artificial intelligence application in medicine (Topol, 2019; Ghassemi et al., 2019). The presented approach demonstrated the reliability of an SVM inspired framework, encompassing the automatic selection of a subset of CTG-derived features, a satisfactory classification performance in terms of AUC, SE, and SP in both the training and validation sets, and interpretable set results suitable to be translatable in the clinical environment. Findings reported in this investigation support the importance of multivariate approaches to investigate the variety of implications resulting from a pathological condition such as late IUGR.

Despite satisfactory and promising classification performance, improvements may be envisioned under various aspects. First, the inclusion of additional features such as the ones inspired to multiscale and fractal analysis might further contribute to classification accuracy as reported in the context of intrapartum (Spilka et al., 2017). Second, the performance of different machine learning approaches as well as deep learning methodologies should be investigated and compared to RFB-SVM-RFE. Lastly, the validation of the presented approach should be carried out on external datasets to ultimately test the model performance as a function of different reference values of the input features. Additionally, it would be relevant to evaluate the validity of the proposed model in the context of the early insurgence of the pathology. A recent dataset of FHR indices extracted from a population of early IUGR fetuses can be found in Signorini et al. (2020a).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: IEEE DataPort—https://ieee-dataport.org/open-access/fetal-heart-rate-features-healthy-and-late-iugr-fetuses.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Azienda Ospedaliera Universitaria Federico II, University of Naples Federico II. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

NP, GM, MC, and MS contributed to the conception and design of the study. GE, ST, and MC carried out the recording of the data. NP, GM, and MS planned and carried out the simulations and analysis of data and took the lead in writing the manuscript. All authors contributed to the interpretation of the results. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

FUNDING

This work was supported by PRIN 2017 MIUR, grant No. ICT4MOMs #2017RR5EW3.

REFERENCES

Arduini, D., Rizzo, G., Piana, A., Brambilla, B. P, and Romanini, C. (1993). Computerized analysis of fetal heart rate—Part I: description of the system (2CTG). *J. Matern. Fetal Invest.* 3, 159–164.

Bauer, A., Kantelhardt, J. W, Bunde, A., Barthel, P., Schneider, R., Malik, M., et al. (2006). Phase-rectified signal averaging detects quasi-periodicities in non-stationary data. *Phys. Stat. Mech. Appl.* 364, 423–434. doi:10.1016/j.physa.2005.08.080

Chen, H. Y., Chauhan, S. P., Ananth, C. V., Vintzileos, A. M., and Abuhamad, A. Z. (2011). Electronic fetal heart rate monitoring and its relationship to neonatal and infant mortality in the United States. *Am. J. Obstet. Gynecol.* 204 (6), 491–e1. doi:10.1016/j.ajog.2011.04.024

Cortes, C., and Vapnik, V. (1995). "Support-vector networks," machine learning, *Springer Science and Business Media LLC.* 20 (3) 273–297. doi:10.1007/bf00994018

Esposito, F. G., Tagliaferri, S., Giudicepietro, A., Giuliano, N., Maruotti, G. M., Saccone, G., et al. (2019). Fetal heart rate monitoring and neonatal outcome in a population of early- and late-onset intrauterine growth restriction. *J. Obstet. Gynaecol. Res.* 45 (7), 1343–1351. doi:10.1111/jog.13981

Faes, L., Marinazzo, D., Jurysta, F., and Nollo, G. (2015). Linear and non-linear brain-heart and brain-brain interactions during sleep. *Physiol. Meas.* 36 (4), 683–698. doi:10.1088/0967-3334/36/4/683

Fanelli, A., Magenes, G., Campanile, M., and Signorini, M. G. (2013). Quantitative assessment of fetal well-being through ctg recordings: a new parameter based on phase-rectified signal average. *IEEE J. Bio. Heal. Inform.* 17 (5), 959–966. doi:10.1109/JBHI.2013.2268423

Ferrario, M., Signorini, M. G., and Magenes, G. (2007). Comparison between fetal heart rate standard parameters and complexity indexes for the identification of severe intrauterine growth restriction. *Methods Inf. Med.* 46 (2), 186–190. doi:10.1055/s-0038-1625404

FIGO (1986). Guidelines for the use of fetal monitoring. *Int. J. Gynecol. Obstet.* 25, 159–167.

Fluss, R., Faraggi, D., and Reiser, B. (2005). Estimation of the Youden index and its associated cutoff point. *Biom. J.* 47 (4), 458–472. doi:10.1002/bimj.200410135

Ghassemi, M., Naumann, T., Schulam, P., Beam, A. L., Chen, I. Y., and Ranganath, R. (2019). Practical guidance on artificial intelligence for health-care data. *Lancet Digit. Health.* 1 (4), e157–e159. doi:10.1016/s2589-7500(19)30084-6

Giuliano, N., Annunziata, M. L., Esposito, F. G., Tagliaferri, S., Di Lieto, A., Magenes, G., et al. (2017). Computerised analysis of antepartum foetal heart parameters: new reference ranges. *J. Obstet. Gynaecol.* 37 (3), 296–304. doi:10.1080/01443615.2016.1239069

Gonçalves, H., Amorim-Costa, C., Ayres-de-Campos, D., and Bernardes, J. (2018). Evolution of linear and nonlinear fetal heart rate indices throughout pregnancy in appropriate, small for gestational age and preterm fetuses: a cohort study. *Comput. Methods Progr. Biomed.* 153, 191–199. doi:10.1016/j.cmpb.2017.10.015

Gonçalves, H., Amorim-Costa, C., Ayres-de-Campos, D., and Bernardes, J. (2017). "Gender-specific evolution of fetal heart rate variability throughout gestation: a study of 8823 cases," in *Early human development*. (North Wall Quay, Dublin: Elsevier Ireland Ltd), Vol. 115, 38–45. doi:10.1016/j.earlhumdev.2017.09.002

Gordijn, S. J., Beune, I. M., Thilaganathan, B., Papageorghiou, A., Baschat, A. A., Baker, P. N., et al. (2016). Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet. Gynecol* 48.(3), 333–339. doi:10.1002/uog.15884

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46 (1–3), 389–422. doi:10.1023/A:1012487302797

Hsu, C.-W., Chang, C.-C., and Lin, C.-J. (2003). *A practical guide to support vector classification*. Taipei, Taiwan: National Taiwan University. doi:10.1007/s11119-014-9370-9

Huhn, E. A., Lobmaier, S., Fischer, T., Schneider, R., Bauer, A., Schneider, K. T., et al. (2011). New computerized fetal heart rate analysis for surveillance of intrauterine growth restriction. *Prenat. Diagn.* 31 (5), 509–514. doi:10.1002/pd.2728

Kohavi, R., and John, G. H. (1997). Wrappers for feature subset selection. *Artif. Intell.* 97 (1–2), 273–324. doi:10.1016/s0004-3702(97)00043-x

Lake, D. E., Richman, J. S., Griffin, M. P., and Moorman, J. R. (2002). Sample entropy analysis of neonatal heart rate variability. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 283 (3), R789–R797. doi:10.1152/ajpregu.00069.2002

Lawson, G. W., Belcher, R., Dawes, G. S., and Redman, C. W. (1983). A comparison of ultrasound (with autocorrelation) and direct electrocardiogram fetal heart rate detector systems. *Am. J. Obstet. Gynecol.* 147 (6), 721–722. doi:10.1016/0002-9378(83)90460-X

Lempel, A., and Ziv, J. (1976). On the complexity of finite sequences. *IEEE Trans. Inf. Theor.* 22 (1), 75–81. doi:10.1109/TIT.1976.1055501

Lin, H., and Lin, C. (2003). A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. *Neural Comput.* 3 (1–32). 16,, 2003 10.1.1.14.6709

Liu, Q., Chen, C., Zhang, Y., and Hu, Z. (2011). Feature selection for support vector machines with RBF kernel. *Artif. Intell. Rev.* 36 (2), 99–115. doi:10.1007/s10462-011-9205-2

Magenes, G., Signorini, M. G, Ferrario, M., and Lunghi, F. (2007). "2CTG2: a new system for the antepartum analysis of fetal heart rate". *11th mediterranean conference on medical and biomedical engineering and computing 2007*. (Berlin, Heidelberg: Springer Berlin Heidelberg), 781–784. doi:10.1007/978-3-540-73044-6_203

Mantel, R., van Geijn, H. P., Caron, F. J., Swartjes, J. M., van Woerden, E. E., and Jongsma, H. W. (1990). Computer analysis of antepartum fetal heart rate: 1. Baseline determination. *Int. J. Bio Med. Comput.* 25 (4), 261–272. doi:10.1016/0020-7101(90)90030-x Available at: http://www.ncbi.nlm.nih.gov/pubmed/2194979 (Accessed March 26, 2018).

Mureşan, D., Rotar, I. C., and Stamatian, F. (2016). The usefulness of fetal Doppler evaluation in early versus late onset intrauterine growth restriction. *Review of the literature. Med. Ultrason.* 18 (1), 103–109. doi:10.11152/mu.2013.2066.181.dop

Parra-Saavedra, M., Crovetto, F., Triunfo, S., Savchev, S., Peguero, A., Nadal, A., et al. (2013). Placental findings in late-onset SGA births without Doppler signs

of placental insufficiency. *Placenta.* 34 (12), 1136–1141. doi:10.1016/j.placenta.2013.09.018

Pincus, S. M., and Viscarello, R. R. (1992). Approximate entropy: a regularity measure for fetal heart rate analysis. *Obstet. Gynecol.* 79 (2), 249–255.

Poudel, R., McMillen, I. C., Dunn, S. L., Zhang, S., and Morrison, J. L. (2015). Impact of chronic hypoxemia on blood flow to the brain, heart, and adrenal gland in the late-gestation IUGR sheep fetus. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 308 (3), R151–R162. doi:10.1152/ajpregu.00036.2014

Rabinowitz, R., Persitz, E., and Sadovsky, E. (1983). The relation between fetal heart rate accelerations and fetal movements. *Obstet. Gynecol.* 61 (1), 16–18.

Rivolta, M. W., Stampalija, T., Frasch, M. G., and Sassi, R. (2019). Theoretical value of deceleration capacity points to deceleration Reserve of fetal heart rate. *IEEE Trans. Bio. Eng.* 2019, 1. doi:10.1109/tbme.2019.2932808

Rosenberg, A. (2008). The IUGR newborn. *Semin Perinatol.* 32 (3), 219–224. doi:10.1053/j.semperi.2007.11.003

Sanz-Cortés, M., Carbajo, R. J., Crispi, F., Figueras, F., Pineda-Lucena, A., and Gratacós, E. (2013). Metabolomic profile of umbilical cord blood plasma from early and late intrauterine growth restricted (IUGR) neonates with and without signs of brain vasodilation. *PLoS One* 8 (12), e80121. doi:10.1371/journal.pone.0080121

Sharma, D., Shastri, S., and Sharma, P. (2016). Intrauterine growth restriction: antenatal and postnatal aspects. *Clin. Med. Insights Pediatr.* 10, 67. doi:10.4137/cmped.s40070

Signorini, M. G., Magenes, G., Cerutti, S., and Arduini, D. (2003). Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocography recordings. *IEEE Trans. Biomed. Eng.* 50 (3), 365–374. doi:10.1109/TBME.2003.808824

Signorini, M. G., Pini, N., Malovini, A., Bellazzi, R., and Magenes, G. (2020a). Dataset on linear and non-linear indices for discriminating healthy and IUGR fetuses. *Data in Brief Elsevier BV.* 29. 105164. doi:10.1016/j.dib.2020.105164

Signorini, M. G., Pini, N., Malovini, A., Bellazzi, R., and Magenes, G. (2020b). Integrating machine learning techniques and physiology based heart rate features for antepartum fetal monitoring. *Comput. Methods Progr. Biomed.* 185, 105015. doi:10.1016/j.cmpb.2019.105015

Smith, J., Murphy, M., and Kandasamy, Y. (2013). The IUGR infant: a case study and associated problems with IUGR infants. *J. Neonatal Nurs.* 19 (2), 46–53. doi:10.1016/j.jnn.2012.12.005

Spilka, J., Frecon, J., Leonarduzzi, R., Pustelnik, N., Abry, P., and Doret, M. (2017). Sparse support vector machine for intrapartum fetal heart rate classification. *IEEE J. Biomed. Health Inform.* 21 (3), 664–671. doi:10.1109/JBHI.2016.2546312

Spilka, J., Chudáček, V., Koucký, M., Lhotská, L., Huptych, M., Janků, P., et al. (2012). Using nonlinear features for fetal heart rate classification. *Biomed. Signal Process Contr.* 7 (4), 350–357. doi:10.1016/j.bspc.2011.06.008

Stampalija, T., Casati, D., Montico, M., Sassi, R., Rivolta, M. W., Maggi, V., et al. (2015). Parameters influence on acceleration and deceleration capacity based on trans-abdominal ECG in early fetal growth restriction at different gestational age epochs. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 188, 104–112. doi:10.1016/j.ejogrb.2015.03.003

Sun, Y. (2007). Iterative RELIEF for feature weighting: algorithms, theories, and applications. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (6), 1035–1051. doi:10.1109/TPAMI.2007.1093

Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Stand. Meas. Phys. Inter. Clini. Use'Circ.* 93, 1043–1065.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* 25 (1), 44–56. doi:10.1038/s41591-018-0300-7

Villar, J., Papageorghiou, A. T., Pang, R., Ohuma, E. O., Cheikh Ismail, L., Barros, F. C., et al. (2014). The likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st project: the fetal growth longitudinal study and newborn cross-sectional study. *Lancet Diabetes Endocrinol.* 2 (10), 781–792. doi:10.1016/S2213-8587(14)70121-4

Warland, J., and Mitchell, E. A. (2014). A triple risk model for unexplained late stillbirth. *BMC Pregnancy Childbirth* 14 (1), 142. doi:10.1186/1471-2393-14-142

Zhang, D. D., Zhou, X. H., Freeman, D. H., and Freeman, J. L. (2002). A non-parametric method for the comparison of partial areas under ROC curves and its application to large health care data sets. *Stat. Med.* 21 (5), 701–715. doi:10.1002/sim.1011

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 05 May 2021
doi: 10.3389/fped.2021.593889

Fetal Cardiovascular Decompensation During Labor Predicted From the Individual Heart Rate Tracing: A Machine Learning Approach in Near-Term Fetal Sheep Model

Nathan Gold[1,2], *Christophe L. Herry*[3], *Xiaogang Wang*[1,4] *and Martin G. Frasch*[5*]

[1] Department of Mathematics and Statistics, York University, Toronto, ON, Canada, [2] Centre for Quantitative Analysis and Modelling, Fields Institute for Research in Mathematical Science, Toronto, ON, Canada, [3] Dynamical Analysis Laboratory, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada, [4] Institute of Big Data, Qing Hua University, Beijing, China, [5] Department of Obstetrics and Gynecology and Center on Human Development and Disability, University of Washington, Seattle, WA, United States

OPEN ACCESS

Edited by:
Arjan Te Pas,
Leiden University, Netherlands

Reviewed by:
Hylke Salverda,
Leiden University Medical Center, Netherlands
Steffen Pauws,
Tilburg University, Netherlands

***Correspondence:**
Martin G. Frasch
mfrasch@uw.edu

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: 11 August 2020
Accepted: 31 March 2021
Published: 05 May 2021

Citation:
Gold N, Herry CL, Wang X and Frasch MG (2021) Fetal Cardiovascular Decompensation During Labor Predicted From the Individual Heart Rate Tracing: A Machine Learning Approach in Near-Term Fetal Sheep Model. Front. Pediatr. 9:593889. doi: 10.3389/fped.2021.593889

Background: When exposed to repetitive umbilical cord occlusions (UCO) with worsening acidemia, fetuses eventually develop cardiovascular decompensation manifesting as pathological hypotensive arterial blood pressure (ABP) responses to fetal heart rate (FHR) decelerations. Failure to maintain cardiac output during labor is a key event leading up to brain injury. We reported that the timing of the event when a fetus begins to exhibit this cardiovascular phenotype is highly individual and was impossible to predict.

Objective: We hypothesized that this phenotype would be reflected in the individual behavior of heart rate variability (HRV) as measured by root mean square of successive differences of R-R intervals (RMSSD), a measure of vagal modulation of HRV, which is known to increase with worsening acidemia. This is clinically relevant because HRV can be computed in real-time intrapartum. Consequently, we aimed to predict the individual timing of the event when a hypotensive ABP pattern would emerge in a fetus from a series of continuous RMSSD data.

Study Design: Fourteen near-term fetal sheep were chronically instrumented with vascular catheters to record fetal arterial blood pressure, umbilical cord occluder to mimic uterine contractions occurring during human labor and ECG electrodes to compute the ECG-derived HRV measure RMSSD. All animals were studied over a ∼6 h period. After a 1–2 h baseline control period, the animals underwent mild, moderate, and severe series of repetitive UCO. We applied the recently developed machine learning algorithm to detect physiologically meaningful changes in RMSSD dynamics with worsening acidemia and hypotensive responses to FHR decelerations. To mimic clinical scenarios using an ultrasound-based 4 Hz FHR sampling rate, we recomputed RMSSD from FHR sampled at 4 Hz and compared the performance of our algorithm under both conditions (1,000 Hz vs. 4 Hz).

Results: The RMSSD values were highly non-stationary, with four different regimes and three regime changes, corresponding to a baseline period followed by mild, moderate, and severe UCO series. Each time series was characterized by seemingly randomly occurring (in terms of timing of the individual onset) increase in RMSSD values at different time points during the moderate UCO series and at the start of the severe UCO series. This event manifested as an increasing trend in RMSSD values, which counter-intuitively emerged as a period of relative stationarity for the time series. Our algorithm identified these change points as the individual time points of cardiovascular decompensation with 92% sensitivity, 86% accuracy and 92% precision which corresponded to 14 ± 21 min before the visual identification. In the 4 Hz RMSSD time series, the algorithm detected the event with 3 times earlier detection times than at 1,000 Hz, i.e., producing false positive alarms with 50% sensitivity, 21% accuracy, and 27% precision. We identified the overestimation of baseline FHR variability by RMSSD at a 4 Hz sampling rate to be the cause of this phenomenon.

Conclusions: The key finding is demonstration of FHR monitoring to detect fetal cardiovascular decompensation during labor. This validates the hypothesis that our HRV-based algorithm identifies individual time points of ABP responses to UCO with worsening acidemia by extracting change point information from the physiologically related fluctuations in the RMSSD signal. This performance depends on the acquisition accuracy of beat to beat fluctuations achieved in trans-abdominal ECG devices and fails at the sampling rate used clinically in ultrasound-based systems. This has implications for implementing such an approach in clinical practice.

Keywords: HRV, hypotension, brain Injury, Bezold Jarisch reflex, machine learning, time series, anomaly detection, changepoint detection

A. **Why was the study conducted?**

During labor, fetuses may develop pathologically hypotensive arterial blood pressure responses to fetal heart rate (FHR) decelerations triggered by uterine contractions. The timing of this event is difficult to predict clinically. We developed a machine learning method to detect this event from an individual FHR tracing.

B. **What are the key findings?**

This real-time algorithm performs well on noisy FHR data requiring ∼2 hours to train on the individual FHR tracings in the first stage of labor; once trained, the algorithm predicts the event with 92% sensitivity, 86% accuracy, and 92% precision.

The algorithm's performance deteriorates to 50% sensitivity, 21% accuracy, and 27% precision when the FHR is acquired at a sampling rate of 4 Hz used in the ultrasound (CTG) monitors compared to the ECG-derived signal as it can be acquired from maternal abdominal ECG.

C. **What does this study add to what is already known?**

This is the first demonstration of the ability to detect fetal cardiovascular decompensation, a prequel to brain injury, intrapartum. The approach is ready for clinical testing. Computerized CTG monitoring cannot predict fetal acidemia intrapartum as well as ECG-based FHR monitoring. This study adds to this knowledge that a computerized approach for objective detection of cardiovascular compromise from FHR in real-time from an individual FHR tracing also performs better when using ECG-derived FHR tracing than CTG tracing.

INTRODUCTION

Electronic fetal monitoring (EFM) cannot identify fetuses at risk of incipient brain injury. The efforts to identify intrapartum acidemia using EFM have failed, in particular using fetal heart rate (FHR) monitoring, because fetal brain injury is poorly correlated with acidemia (1). Brain compromise due to hypoxia-ischemia (HI) can ensue when the fetal cerebral blood flow is persistently reduced e.g., due to precipitous drop in cerebral perfusion pressure resulting from cardiovascular decompensation (2, 3). Bezold-Jarisch reflex (BJR) is a vagal depressor reflex observed in fetal sheep under the conditions of umbilical cord occlusions (UCO) with worsening acidemia which leads to cardiovascular decompensation (4). We asked whether FHR monitoring can capture the BJR-mediated vagal sensing of acidemia. We studied the relationship between fetal systemic arterial blood pressure (ABP) and FHR in an animal model of human labor.

We had reported that sheep fetuses have an individual cardiovascular phenotype in their responses to increasing acidemia due to repetitive intermittent hypoxia (3). We

hypothesized that such phenotype would be reflected in individual responses of heart rate variability (HRV) as measured by root mean square of successive differences of R-R intervals (RMSSD), a measure of vagal modulation of HRV known to increase with worsening acidemia (5–7). Consequently, a series of continuously computed RMSSD data will consistently predict the event when a hypotensive ABP pattern emerges in an individual fetus (3).

The current standard of EFM relies predominantly on ultrasound-based FHR monitoring. Because the vagally mediated HRV is found on a time scale that is not captured at 4 Hz sampling rate, we also tested the impact of its inherently lower FHR sampling rate precision of 4 Hz vs. the golden standard electrocardiogram (ECG)—derived 1,000 Hz on the ability to individually predict cardiovascular decompensation. We hypothesized that the lower temporal precision will result in a poorer prediction of the timing of cardiovascular decompensation.

MATERIALS AND METHODS

Experimental methods and data acquisition have been presented elsewhere (8). Briefly, fourteen near-term fetal sheep were chronically instrumented with vascular catheters to record fetal arterial blood pressure, umbilical cord occluder to mimic uterine contractions occurring during human labor and ECG electrodes to compute ECG-derived HRV measure RMSSD. Animal care followed the guidelines of the Canadian Council on Animal Care and was approved by the University of Western Ontario Council on Animal Care.

Surgical Preparation

Fourteen near-term ovine fetuses [123 ± 2 days gestational age (GA), term = 145 days] of the mixed breed were surgically instrumented. The anesthetic and surgical procedures and postoperative care of the animals have been previously described (3, 9). Briefly, polyvinyl catheters were placed in the right and left brachiocephalic arteries, the cephalic vein, and the amniotic cavity. Stainless steel electrodes were sewn onto the fetal chest to monitor the electrocardiogram (ECG). A polyvinyl catheter was also placed in the maternal femoral vein. Stainless steel electrodes were additionally implanted biparietally on the dura for the recording of electrocorticogram, ECOG, as a measure of summated brain electrical activity [results reported elsewhere (3, 8, 10)]. An inflatable silicon rubber cuff (*In vivo* Metric, Healdsburg, CA) for UCO induction was placed around the proximal portion of the umbilical cord and secured to the abdominal skin. Once the fetus was returned to the uterus, a catheter was placed in the amniotic fluid cavity. Antibiotics were administered intravenously to the mother (0.2 g of trimethoprim and 1.2 g sulfadoxine, Schering Canada Inc., Pointe-Claire, Canada) and fetus and into the amniotic cavity (1 million IU penicillin G sodium, Pharmaceutical Partners of Canada, Richmond Hill, Canada). Amniotic fluid lost during surgery was replaced with warm saline. The uterus and abdominal wall incisions were sutured in layers and the catheters exteriorized through the maternal flank and secured to the back of the ewe in a plastic pouch.

Postoperatively, animals were allowed 4 days to recover prior to experimentation and daily antibiotic administration was continued intravenously to the mother (0.2 g trimethoprim and 1.2 g sulfadoxine), into the fetal vein and the amniotic cavity (1 million IU penicillin G sodium, respectively). Arterial blood was sampled for evaluation of the maternal and fetal condition and catheters were flushed with heparinized saline to maintain patency. Animals were 130 ± 1 day GA on the first day of the experimental study.

Experimental Procedure

All animals were studied over a ∼6 h period (**Figure 1**). Fetal chronic hypoxia was defined as arterial O_2Sat <55% as measured on postoperative days 1–3 and at baseline prior to beginning the UCOs. The first group comprised five fetuses that were also spontaneously hypoxic (n = 5, H/UCO). The second group of fetuses was normoxic (O_2Sat>55% before UCOs) (n = 9, N/UCO). The experimental protocol has been reported (7, 9, 11). After a 1–2 h baseline control period, the animals underwent mild, moderate, and severe series of repetitive UCOs by graduated inflation of the occluder cuff with a saline solution. During the first hour following the baseline period, mild variable FHR decelerations were performed with a partial UCO for 1 min duration every 2.5 min, with the goal of decreasing FHR by ∼30 bpm, corresponding to a ∼50% reduction in umbilical blood flow (12, 13). During the second hour, moderate variable FHR decelerations were performed with increased partial UCO for 1 min duration every 2.5 min with the goal of decreasing FHR by ∼60 bpm, corresponding to a ∼75% reduction in umbilical blood flow (13). Animals underwent severe variable FHR decelerations with complete UCO for 1 min duration every 2.5 min until the targeted fetal arterial pH of <7.0 was detected or 2 h of severe UCO had been carried out, at which point the repetitive UCOs were terminated. These animals were then allowed to recover for 48 h following the last UCO. Fetal arterial blood samples were drawn at baseline, at the end of the first UCO of each series (mild, moderate, severe), and at 20 min intervals (between UCOs) throughout each of the series, as well as at 1, 24, and 48 h of recovery. For each UCO series blood gas sample and the 24 h recovery sample of 0.7 ml of fetal blood was withdrawn, while 4 ml of fetal blood was withdrawn at baseline, at pH_{nadir} < 7.00, and at 1 and 48 h of recovery. The amounts of blood withdrawn were documented for each fetus and replaced with an equivalent volume of maternal blood at the end of day 1 of the study.

All blood samples were analyzed for blood gas values, pH, glucose, and lactate with an ABL-725 blood gas analyzer (Radiometer Medical, Copenhagen, Denmark) with temperature corrected to 39.0°C. Plasma from the 4 ml blood samples was frozen and stored for cytokine analysis, reported elsewhere.

After the 48 h recovery blood sample, the ewe and the fetus were killed by an overdose of barbiturate (30 mg sodium pentobarbital IV, MTC Pharmaceuticals, Cambridge, Canada). A post mortem was carried out during which fetal sex and weight were determined and the location and function of the umbilical occluder were confirmed. The fetal brain was

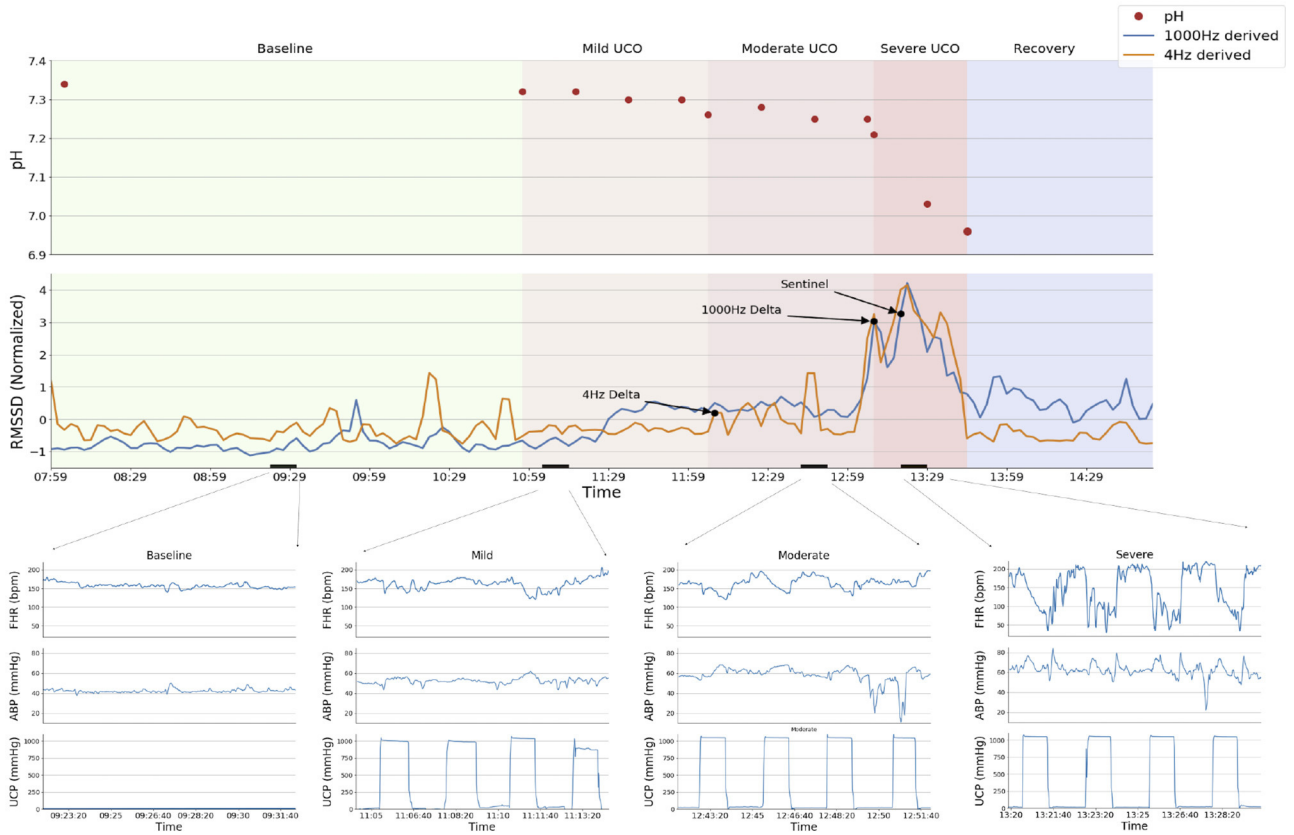

FIGURE 1 | Experimental protocol and a representative example of the analytical approach (animal ID 473378). The RMSSD time series derived from 1,000 Hz (blue) and 4 Hz (orange) sampled ECG are displayed superimposed in the top panel, with declared change points from the BOCPD algorithm and the Sentinel (expert detection) marked with arrows. Sequential fetal arterial pH measurements are indicated. Experimental stages are demarcated by background colors; short black bars over X-axis indicate the zoomed-in segments shown in the bottom panel. Bottom panels show fetal heart rate (FHR, bpm), fetal arterial blood pressure (ABP, mmHg) and umbilical cord pressure (UCP, mmHg) indicating when UCO were triggered (increasing UCP). Note the failure of the change point algorithm to detect the sentinel time point correctly (i.e., around the Sentinel time point) when using 4 Hz—derived RMSSD signal: the detection occurs 1 h earlier than with the 1,000 Hz signal. This is due to unphysiological fluctuations in FHR variability at baseline as demonstrated in Durosier et al. (7) and Li et al. (19).

perfusion-fixed and subsequently dissected and processed for later immunohistochemical study as reported (14).

Data Acquisition and Analysis

A computerized data acquisition system was used to record fetal systemic arterial and amniotic pressures and the ECG signal, as described (7). All signals were monitored continuously throughout the experiment. Arterial and amniotic pressures were measured using Statham pressure transducers (P23 ID; Gould Inc., Oxnard, CA). Arterial blood pressure (ABP) was determined as the difference between instantaneous values of arterial and amniotic pressures. A PowerLab system was used for data acquisition and analysis (Chart 5 For Windows, ADInstruments Pty Ltd, Castile Hill, Australia). Pressures, ECOG and ECG were recorded and digitized at 1,000 Hz for further study. For ECG, a 60 Hz notch filter was applied.

R peaks of ECG were used to derive the heart rate variability (HRV) times series. Beat detection was performed using a mix of two algorithms, a custom wavelet-based detection and Elgendi's method with an added refractory period step (15). Both methods include bandpass filtering as an initial step that removes baseline wandering and high frequency noise. Beat detection was also verified for accuracy using a custom developed ECG annotation and reviewing tool. This was necessary to validate beat detection in UCO periods where the noise level was high and where there were artifacts generated by the contractions. R-R intervals were further filtered based on the morphology of the ECG waveforms, the level of noise/artifacts within short windows and the proportion of disconnected/saturated periods, if any (16, 17). Windows of low quality were not retained in the HRV analysis. Low quality was defined over analysis windows (5 min) as a weighted sum of the percentage of time without non-physiologic beats (artifacts), the percentage of time uninterrupted by disconnections/saturations, the percentage of time with high quality beats according to Clifford et al.'s (16, 17). Ectopies were not filtered out as there were a large number of them

during UCO periods and it would effectively remove most of the UCO periods if filtered out. The average percentage of original ECG signal ultimately discarded for the HRV analysis was 5.5% (range: 1–20%).

The time series of R-R peak intervals were uniformly resampled at 4 Hz. Technically, the resampling was performed as an interpolation since we need to go from a pseudo-frequency of 2.5–3 Hz to a sampling frequency of 4 Hz. The interpolation method used was a piecewise cubic Hermite interpolation. Next, the RMSSD was calculated continuously on both the original R-R interval time series (with 1-millisecond resolution) and the R-R interval time series resampled at 4 Hz, from each 5 min HRV segment in 2.5 min overlapping sliding windows. For the ∼6-h time series, this corresponded to roughly 150 data points.

During UCO series, the point at which hypotensive ABP responses to UCO had been detected by "expert" visual detection was termed ABP "sentinel," defined as the time between the onset of such ABP responses to UCO and the time when pH nadir (pH < 7.00) is reached in each fetus.

To detect changes in RMSSD values corresponding to the above sentinel time point in the ABP responses, we used the previously reported machine learning algorithm, referred to as Delta point method, based on change point detection (18). Briefly, Delta point method is a real-time change point detection method, robust to false-alarms, designed to filter a vector of suspected change points. It proceeds by fitting a probabilistic Gaussian process model to the RMSSD time series baseline data and computing online predictions of the RMSSD values within the range of the model. Suspected change-points are declared as significant ($p < 0.05$) deviations from pointwise model predictions and observations. These are viewed as observations of a doubly stochastic Poisson process, with observation rate governed by the Gaussian process model. Based upon this theory, the points are grouped into time intervals, within which the Delta point is selected as the most significant change.

To perform hyper-parameter training, we segmented the data into a 60 point training set, or 2.5-h training time on the baseline and mild UCO periods of each time series (i.e., corresponding to the first stage of labor). Our method uses an $n = 10$ point or 25-min interval to segment the time series for delta point evaluation. The choice of 10 points or 25-min interval is to provide a reasonable number of points per interval for the Delta point method, so that a reasonable average may be calculated for the average run in each interval.

We defined a successful detection as the agreement between the Delta point and the sentinel value, with Delta point detection no later than 2 min behind expert detection. False negative detections were defined as Delta point being declared 2 min behind expert detection. False positive detections were defined as detection occuring 25 min prior to expert detection, corresponding to one Delta point sampling interval. This demonstrates the effectiveness of the method, suggesting clinical benefits for earlier decision making.

Statistical Analysis

The differences in the change point detection at 4 Hz compared to 1,000 Hz were evaluated with the Wilcoxon signed-rank test with a $P < 0.05$ was considered significant. Detection performance was analyzed by computing the accuracy, sensitivity, and precision of the method defined as,

$$\text{Accuracy} = \frac{\text{Successful detections}}{\text{Number of examples}} \quad (1)$$

$$\textbf{Sensitivity} = \frac{\textbf{True Positive}}{\textbf{True Positive} + \textbf{False Negative}} \quad (2)$$

$$\textbf{Precision} = \frac{\textbf{True Positive}}{\textbf{True Positive} + \textbf{False Positive}} \quad (3)$$

TABLE 1 | Confusion matrix.

1,000 Hz	Positive	Negative	4 Hz	Positive	Negative
Positive	12	1	**Positive**	3	8
Negative	1	0	**Negative**	3	0

Results

The physiological characteristics of the experimental groups have been reported (8, 10, 11).

Delta point method was able to match the expert prediction with Delta point declaration occurring at a median 8.5 (IQR = 10.5) minutes before ABP sentinel time. This corresponded to 92% sensitivity, 86% accuracy, and 92% precision.

In the 4 Hz RMSSD time series, the algorithm triggered change point at a median 36 (IQR = 44.3) minutes failing to match the expert prediction by yielding 8 times earlier detection times than at 1,000 Hz, i.e., producing false positive alarms in 8 out of 14 cases ($p = 0.003$). This corresponded to 50% sensitivity, 21% accuracy, and 27% precision. We report the confusion matrix for both the 1,000 Hz RMSSD and 4 HZ RMSSD time series in **Table 1**.

A representative example of the experimental data is shown in **Figure 1** and the individual findings for all subjects are reported in **Table 2**.

The visual inspection of the RMSSD tracings suggested that the overestimation of the baseline FHR variability by RMSSD at the 4 Hz sampling rate is the cause of this false detection phenomenon. To verify this assumption we determined the RMSSD values computed from the 1,000 Hz and 4 Hz sampled FHR data sets at baseline and during the UCO series. Confirming our hypothesis, we found a smaller difference in the average normalized RMSSD values during the UCO series compared to the baseline in the 4 Hz data set (0.52 ± 0.16) compared to the 1,000 Hz data set (0.85 ± 0.4, $p = 0.027$).

DISCUSSION

Principal Findings

Our findings validate the hypothesis that Delta point method, applied to the FHR-derived HRV measure RMSSD, identifies individual time points of ABP responses to UCO with worsening acidemia by extracting change point information from the physiologically related fluctuations in RMSSD time series. The present findings also show the dependence of this method on high temporal precision of FHR acquisition to capture correctly

TABLE 2 | Performance of the anomaly detection algorithm in predicting the individual time points of cardiovascular decompensation from FHR.

Group	Animal	Sentinel	1,000 Hz detection	4 Hz detection	1,000 Hz delta	4 Hz delta
H_UCO	8003	15:56:00	15:49:00	15:52:00	0:07	0:04
H_UCO	473351	13:38:00	13:28:00	13:04:00	0:10	0:34
H_UCO	473362	11:05:00	11:03:00	11:35:00	0:02	0:30
H_UCO	473376	12:36:00	12:38:00	11:59:00	0:02	0:37
H_UCO	473726	12:04:00	11:50:00	11:54:00	0:14	0:10
N_UCO	461060	12:42:00	12:31:00	12:21:00	0:11	0:21
N_UCO	473361	12:51:00	12:36:00	12:16:00	0:15	0:35
N_UCO	473352	13:17:00	12:53:00	12:06:00	0:24	1:11
N_UCO	473377	12:12:00	12:14:00	12:50:00	0:02	0:38
N_UCO	473378	13:22:00	13:09:00	12:09:00	0:13	1:13
N_UCO	473727	11:03:00	11:10:00	11:08:00	0:07	0:05
N_UCO	5054	12:53:00	11:27:00	11:19:00	1:26	1:34
N_UCO	5060	11:26:00	11:24:00	10:29:00	0:02	0:57
N_UCO	473360	13:59:00	13:52:00	11:55:00	0:07	2:04

Sentinel, time of detecting the onset of pathological ABP decreases during UCOs by an expert (visual analysis); 1,000 and 4 Hz detection, times of detecting the same using the change point algorithm on RMSSD data derived from 1,000 or 4 Hz sampled ECG; 1,000 delta and 4 Hz delta, the time difference (sentinel-1,000 or sentinel-4 Hz) between expert and change point algorithm detection performance: detection by the algorithm preceded in most cases the expert detection, median 8.5 (IQR = 10.5) minutes and median 36 (IQR = 44.3) minutes, respectively; Red font, cases when the algorithm detection happened after the expert detection; note that in the case of 4 Hz delta, 2 out of 3 instances the detection was more than 30min too late compared to ∼3min too late in the three cases at 1,000 Hz.

the physiological fluctuations of FHR at baseline. This is in line with the previous observations in the pregnant sheep model and human fetuses intrapartum (7, 19).

RESULTS

We had reported consistent changes in fetal brain electrical activity, the electrocorticogram (ECOG), with amplitude suppression and frequency increase during FHR decelerations accompanied by highly correlated pathological decreases in fetal ABP, referred to as adaptive brain shutdown (3). These changes in ECOG occurred on average 50 min prior to attaining a severe degree of acidemia (i.e., fetal arterial pH<7.00). However, we noted a high degree of inter-individual variability in the timing of the onset of these brain electrical and cardiovascular responses. Importantly for the neonatal outcome, we found a relationship between the ensuing neuroinflammation measured by the number of microglia, the brain's immune cells, and the timing of the adaptive brain shutdown onset (14). An individualized and timely detection of the onset of hypotensive responses to worsening acidemia and hence the timing of the adaptive brain shutdown would provide clinically relevant information on the degree of neuroinflammation after birth. Perinatal neuroinflammation has been identified as relevant prognostically not only short-term during early life, but also long-term for adult neurodevelopmental sequelae (20–29).

We suggest that the robust performance of the algorithm is owed to selecting causally linked phenomena which are reflected in the two different time series: RMSSD is known to rise with worsening acidemia due to chemoreceptors activation for example (6, 7). Meanwhile, fetal ABP responses to worsening acidemia deteriorate over time with an initial phase of hypertensive responses during each UCO to compensate for the drop in FHR, followed by the gradual decline of this hypertensive component and eventually ensuing pathological hypotensive ABP responses (3). This is at least partially due to a cardiac decompensation with growing levels of acidemia (30, 31). Acidemia impacts myocardial contractility which decreases cardiac output and ABP. It is plausible to expect that such transition in cardiac behavior will be reflected in HRV, RMSSD in particular, because HRV reflects not only the influences of the autonomic nervous system's vagal modulation of the cardiac sinus node activity, it also depends on the intrinsic cardiac rhythm fluctuations and health as evidenced by a decrease in HRV in patients after heart transplants and by presence of intrinsic HRV as early as in term fetuses of gestational age similar to the present study (31–35).

The RMSSD time series were highly non-stationary, with four different regimes and three regime changes, corresponding to a baseline period followed by mild, moderate and severe UCO series. Each time series was characterized by seemingly randomly occurring (in terms of timing of the individual onset) increase in RMSSD values at different time points during the moderate UCO series and at the start of the severe UCO series. This event manifested as an increasing trend in RMSSD values, which counter-intuitively emerged as a period of relative stationarity for the time series. The Delta point algorithm effectively declared these points as the change point of clinical importance. Overall, we found the Delta point algorithm's predictions to be reliable even in the instances when the signals were noisy (18). This is based on the tests of the algorithm in various data sets as published (18) and on our observation that here, to mimic the online recording situation, no correction for ectopies or

non-sinus rhythms was undertaken on RMSSD as is usually done for HRV offline processing (34). To our knowledge, no comparable statistical or machine learning methods for FHR analysis exist.

The reliance on a high-quality RMSSD signal (i.e., derived from 1,000 Hz sampled true beat-to-beat variability signal) is also what explains the failure of the algorithm to detect relevant changepoints at 4 Hz sampling rate when the RMSSD signal becomes distorted due to undersampling and the resulting overestimation of baseline variability (7, 19).

Clinical Implications

We demonstrate that computerized FHR monitoring intrapartum deploying machine learning can detect fetal cardiovascular decompensation during labor. Considering the average duration of labor of 12 h for nullipara, the requirement of a 2-h training window on the individual patient's data for the proposed algorithm is trivial (36, 37). Possible decision support such an algorithm can provide is alerting the healthcare provider to ease on contractions or to expedite the delivery to prevent fetal brain injury. Development of the actual clinical workflow will require retrospective and prospective clinical studies.

Our findings have direct clinical implications since high precision HRV can be recorded non-invasively in human fetuses from maternal abdominal ECG (38–42). Moreover, the present results validate and extend the insight we and others reported earlier in sheep and human fetuses whereby the reduced sampling rate of FHR acquisition decreases the precision of HRV—derived measures such as RMSSD for the detection of acidemia (7, 19, 43). Here, we show that the Delta point method performs 3-times more precisely in alerting to fetal cardiovascular decompensation when the underlying FHR signal was sampled at the gold standard 1,000 Hz rate available with today's fetal ECG monitors rather than at the 4 Hz rate as acquired with the ultrasound monitors.

Research Implications

Future prospective clinical studies will investigate the utility of this discovery in the early detection of fetal cardiovascular compromise intrapartum using EFM. Our findings indicate the superiority of abdominal ECG-derived FHR signal for the prediction of cardiovascular decompensation. The present machine learning approach relies on the individual tracing to learn its properties and detect the timing of fetal cardiovascular compromise. That is, unlike most of the artificial intelligence (AI) technologies based on other machine learning methodologies or deep learning (artificial neural networks), our algorithm does not require a large amount of data from multiple subjects (thousands of subjects) to be fed into it in order to perform. Nevertheless, the advent of deep learning may also open new applications for more precise, individualized decision support using the conventional ultrasound-derived FHR tracings. In this context, future studies could focus on building big datasets of FHR recordings intrapartum to enable large scale testing of the AI-based algorithms such as the one presented here or the ones based on deep learning approaches, e.g., as recently pioneered in EFM by Georgieva et al. (44).

Strengths and Limitations

The present findings from an established preclinical translational experiment present a conceptual advance in the clinical EFM demonstrating a novel machine learning approach for individualized detection of fetal cardiovascular compromise using FHR. The individual machine learning time of ∼2 h during the first stage of labor is clinically realistic. The main limitation of the present study is that its insights are derived from an animal model paradigm, albeit well validated. As such, prospective human clinical studies of FHR intrapartum are needed. Such clinical studies will also shed light on our *a priori* choice of 25 min prior to the sentinel event as a cut-off for true positive detection. It is possible that an earlier detection and decision support for intervention in labor will be found beneficial for mother's and child's health. In such case, the 4 Hz-based conventional ultrasound FHR monitoring may turn out to also be amenable to such an algorithm. The risk of increasing the already alarmingly high rate of unnecessary cesarean sections speaks against this notion at this time. Furthermore, our approach so far took no advantage of the information contained in the changes in the uterine pressure during contractions in the first and second stages of pushing and the FHR response to it as is routinely done clinically during the FHR assessment. A combination of the present machine learning approach with information from uterine contractions will likely boost the performance of the presented algorithm in a clinical setting.

Conclusions

The novelty of the current work is that its EFM algorithm permits statistical-level predictions about concomitant changes in individual FHR tracings which alert about fetal cardiovascular decompensation, an important mechanistic prequel to brain injury. The presented approach now awaits direct clinical validation in retrospective or prospective clinical studies.

CONDENSATION

Fetal heart rate (FHR) algorithm based on machine learning from individual FHR tracings detects early cardiovascular compromise in a sheep model of human labor.

DATA AVAILABILITY STATEMENT

The original contributions generated for the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The animal study was reviewed and approved by University of Western Ontario Council on Animal Care.

AUTHOR CONTRIBUTIONS

NG, XW, and MF conceived of the manuscript. NG, CH, and MF wrote the initial draft. NG and CH analyzed the data. NG, CH,

XW, and MF contributed to the draft, read and approved the final version of the manuscript. MF holds patents on fetal EEG and ECG monitoring: US9,215,999 and WO2018160890. The authors have declared that no further conflict of interest exists.

FUNDING

MF was funded by the Canadian Institutes for Health Research (CIHR Grant Number: 123489).

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Bryan Richardson and his Perinatal Research Lab at the University of Western Ontario for the original design of the animal experiments that enabled the acquisition of the dataset underlying the present study.

SUPPLEMENTARY METHODS - EXPLANATION OF DELTA POINT METHOD

The Delta point method is a change point detection algorithm for the online interrogation of suspected change points in non-stationary time series. The algorithm proceeds by identifying suspected change points in a non-stationary time series by fitting a non-parametric function representation to the real-time observed time series data. The fitted functional representation then forecasts future values of the time series and a statistical algorithm is applied to the realized and forecasted values to determine if a significant divergence occurs between the respective values. When a statistically significant difference is observed, a suspected change point is declared.

Due to the noisy nature of biologically collected non-stationary time series, this noise often causes many false-positive detected change points for which real-time processing must be applied to determine the temporal location of a true change. The Delta point method uses the expanding returned vector of suspected change points and applies a windowing procedure to the temporal locations of the suspected changes. The rate of suspected change points is then fit to a doubly-stochastic point process to determine the rate at which suspected change points occur in the time series overall, as well as within each window. Due to the functional representation of the time series, this additional information is used to tune the estimated hazard rate of the point process. Based upon this representation, the window of interest is determined as the location with the greatest temporal difference between suspected change points. The Delta point - the change point of true interest - is then selected as the last occurring change point in the identified window.

REFERENCES

1. Cahill AG, Mathur AM, Smyser CD, Mckinstry RC, Roehl KA, López JD, et al. Neurologic injury in acidemic term infants. *Am J Perinatol*. (2017) 34:668–75. doi: 10.1055/s-0036-1597135
2. Astrup J. Energy-requiring cell functions in the ischemic brain. Their critical supply and possible inhibition in protective therapy. *J Neurosurg*. (1982) 56:482–97. doi: 10.3171/jns.1982.56.4.0482
3. 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:e22100. doi: 10.1371/journal.pone.0022100
4. Frasch MG, Mansano R, Ross MG, Gagnon R, Richardson BS. Do repetitive umbilical cord occlusions (UCO) with worsening acidemia induce the Bezold-Jarisch reflex (BJR) in the ovine fetus near term? *Reprod Sci*. (2008) 15:129A.
5. Frasch MG, Mueller T, Wicher C, Weiss C, Loehle M, Schwab K, et al. Fetal body weight and the development of the control of the cardiovascular system in fetal sheep. *J Physiol London*. (2007) 579:893–907. doi: 10.1113/jphysiol.2006.124800
6. Frasch MG, Mueller T, Weiss C, Schwab K, Schubert H, Schwab M. Heart rate variability analysis allows early asphyxia detection in ovine fetus. *Reprod Sci*. (2009) In print.:509–517. doi: 10.1177/1933719108327597
7. Durosier LD, Green G, Batkin I, Seely AJ, Ross MG, Richardson BS, et al. Sampling rate of heart rate variability impacts the ability to detect acidemia in ovine fetuses near-term. *Front Pediatr*. (2014) 2:38. doi: 10.3389/fped.2014.00038
8. 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:e108119. doi: 10.1371/journal.pone.0108119
9. Ross MG, Jessie M, Amaya K, Matushewski B, Durosier LD, Frasch MG, et al. Correlation of arterial fetal base deficit and lactate changes with severity of variable heart rate decelerations in the near-term ovine fetus. *Am J Obstet Gynecol*. (2013) 208:285.e1–6. doi: 10.1016/j.ajog.2012.10.883
10. Frasch MG, Durosier LD, Gold N, Cao M, Matushewski B, Keenliside L, et al. Adaptive shut-down of EEG activity predicts critical acidemia in the near-term ovine fetus. *Physiol Rep*. (2015) 3:e12435. doi: 10.14814/phy2.12435
11. Amaya KE, Matushewski B, Durosier LD, Frasch MG, Richardson BS, Ross MG. Accelerated acidosis in response to variable fetal heart rate decelerations in chronically hypoxic ovine fetuses. *Am J Obstet Gynecol*. (2016) 214:270.e1–270.e8. doi: 10.1016/j.ajog.2015.09.084
12. Itskovitz J, LaGamma EF, Rudolph AM. Heart rate and blood pressure responses to umbilical cord compression in fetal lambs with special reference to the mechanism of variable deceleration. *Am J Obstet Gynecol*. (1983) 147:451–7. doi: 10.1016/S0002-9378(16)32243-8
13. Richardson BS, Rurak D, Patrick JE, Homan J, Carmichael L. Cerebral oxidative metabolism during sustained hypoxaemia in fetal sheep. *J Dev Physiol*. (1989) 11:37–43.
14. Xu A, Durosier LD, Ross MG, Hammond R, Richardson BS, Frasch MG. Adaptive brain shut-down counteracts neuroinflammation in the near-term ovine fetus. *Front Neurol*. (2014) 5:110. doi: 10.3389/fneur.2014.00110
15. Elgendi M. Fast QRS detection with an optimized knowledge-based method: evaluation on 11 standard ECG databases. *PLoS ONE*. (2013) 8:e73557. doi: 10.1371/journal.pone.0073557
16. Clifford GD, Behar J, Li Q, Rezek I. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiol Meas*. (2012) 33:1419–33. doi: 10.1088/0967-3334/33/9/1419
17. Clifford GD, McSharry PE, Tarassenko L. Characterizing artefact in the normal human 24-hour RR time series to aid identification and artificial replication of circadian variations in human beat to beat heart rate using a simple threshold. In: *Computers in Cardiology*. Memphis, TN (2002). p. 129–132.
18. Gold N, Frasch MG, Herry CL, Richardson BS, Wang X. A doubly stochastic change point detection algorithm for noisy biological signals. *Front Physiol*. (2017) 8:1112. doi: 10.3389/fphys.2017.01112
19. Li X, Xu Y, Herry C, Durosier LD, Casati D, Stampalija T, et al. Sampling frequency of fetal heart rate impacts the ability to predict pH and BE at birth: a retrospective multi-cohort study. *Physiol Meas*. (2015) 36:L1–12. doi: 10.1088/0967-3334/36/5/L1

20. Garzoni L, Faure C, Frasch MG. Fetal cholinergic anti-inflammatory pathway and necrotizing enterocolitis: the brain-gut connection begins in utero. *Front Integr Neurosci*. (2013) 7:57. doi: 10.3389/fnint.2013.00057
21. Grether JK, Nelson KB. Maternal infection and cerebral palsy in infants of normal birth weight. *JAMA*. (1997) 278:207–11. doi: 10.1001/jama.278.3.207
22. al-Haddad BJS, Jacobsson B, Chabra S, Modzelewska D, Olson EM, Bernier R, et al. Long-term risk of neuropsychiatric disease after exposure to infection *in utero*. *JAMA Psychiatr*. (2019) 76:594–602. doi: 10.1001/jamapsychiatry.2019.0029
23. Pike K, Brocklehurst P, Jones D, Kenyon S, Salt A, Taylor D, et al. Outcomes at 7 years for babies who developed neonatal necrotising enterocolitis: the ORACLE Children Study. *Arch Dis Child Fetal Neonatal Ed*. (2012) 97:F318–22. doi: 10.1136/fetalneonatal-2011-300244
24. Gotsch F, Romero R, Kusanovic JP, Mazaki-Tovi S, Pineles BL, Erez O, et al. The fetal inflammatory response syndrome. *Clin Obstet Gynecol*. (2007) 50:652–83. doi: 10.1097/GRF.0b013e31811ebef6
25. Spencer SJ, Auer RN, Pittman QJ. Rat neonatal immune challenge alters adult responses to cerebral ischaemia. *J Cereb Blood Flow Metab*. (2006) 26:456–67. doi: 10.1038/sj.jcbfm.9600206
26. Rees S, Inder T. Fetal and neonatal origins of altered brain development. *Early Hum Dev*. (2005) 81:753–61. doi: 10.1016/j.earlhumdev.2005.07.004
27. Bilbo SD. Early-life infection is a vulnerability factor for aging-related glial alterations and cognitive decline. *Neurobiol Learn Memory*. (2010) 94:57–64. doi: 10.1016/j.nlm.2010.04.001
28. Frasch MG, Snir G, Antonelli MC. Autism spectrum disorder: a neuro-immunometabolic hypothesis of the developmental origins. *arXiv*. (2019). Available online at: http://arxiv.org/abs/1909.05198
29. Desplats P, Gutierrez AM, Antonelli MC, Frasch MG. Microglial memory of early life stress and inflammation: susceptibility to neurodegeneration in adulthood. *Neurosci Biobehav Rev*. (2019) 2019:S0149-7634(18)30973-4. doi: 10.1016/j.neubiorev.2019.10.013
30. Yumoto Y, Satoh S, Fujita Y, Koga T, Kinukawa N, Nakano H. Noninvasive measurement of isovolumetric contraction time during hypoxemia and acidemia: fetal lamb validation as an index of cardiac contractility. *Early Hum Dev*. (2005) 81:635–42. doi: 10.1016/j.earlhumdev.2005.04.004
31. Bernardi L, Salvucci F, Suardi R, Solda PL, Calciati A, Perlini S, et al. Evidence for an intrinsic mechanism regulating heart rate variability in the transplanted and the intact heart during submaximal dynamic exercise? *Cardiovasc Res*. (1990) 24:969–81. doi: 10.1093/cvr/24.12.969
32. Yaniv Y, Ahmet I, Liu J, Lyashkov AE, Guiriba TR, Okamoto Y, et al. Synchronization of sinoatrial node pacemaker cell clocks and its autonomic modulation impart complexity to heart beating intervals. *Heart Rhythm*. (2014) 11:1210–9. doi: 10.1016/j.hrthm.2014.03.049
33. Yaniv Y, Lyashkov AE, Lakatta EG. The fractal-like complexity of heart rate variability beyond neurotransmitters and autonomic receptors: signaling intrinsic to sinoatrial node pacemaker cells. *Cardiol Pharmacol*. (2013) 2:111. doi: 10.4172/2329-6607.1000111
34. Seely AJ, Green GC, Bravi A. Continuous multiorgan variability monitoring in critically illpatients–complexity science at the bedside. *Conf Proc IEEE Eng Med Biol Soc*. (2011) 2011:5503–6. doi: 10.1109/IEMBS.2011.6091404
35. Frasch MG, Herry C, Niu Y, Giussani DA. First evidence that intrinsic fetal heart rate variability exists and is affected by chronic hypoxia. *bioRxiv*. (2018) 242107. doi: 10.1101/242107
36. Frasch MG. Letter to the Editor: mind the gap: epistemology of heart rate variability. *Am J Physiol Regul Integr Comp Physiol*. (2020) 319:R343–4. doi: 10.1152/ajpregu.00183.2020
37. Gold N, Frasch M. Mini-commentary on BJOG-20-1394. R2. What's pH got to do with it? *Authorea Preprints*. (2021). doi: 10.22541/au.161215910.00224413/v1
38. Frasch MG, Xu Y, Stampalija T, Durosier LD, Herry C, Wang X, et al. Correlating multidimensional fetal heart rate variability analysis with acid-base balance at birth: a prospective cohort study. *Physiol Meas*. (2014) 35:L1–12. doi: 10.1088/0967-3334/35/12/L1
39. Frasch MG. Saving the brain one heartbeat at a time. *J Physiol*. (2018) 596:5503–4. doi: 10.1113/JP275434
40. 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. doi: 10.3389/fphys.2017.00277
41. 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) 205:47.e1–5. doi: 10.1016/j.ajog.2011.02.066
42. Cohen WR, Ommani S, Hassan S, Mirza FG, Solomon M, Brown R, et al. Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes. *Acta Obstet Gynecol Scand*. (2012) 91:1306–13. doi: 10.1111/j.1600-0412.2012.01533.x
43. Goncalves H, Costa A, Ayres-de-Campos D, Costa-Santos C, Rocha AP, Bernardes J. Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability. *Med Biol Eng Comput*. (2013) 51:665–76. doi: 10.1007/s11517-013-1036-7
44. Petrozziello A, Redman CWG, Papageorghiou AT, Jordanov I, Georgieva A. Multimodal convolutional neural networks to detect fetal compromise during labor and delivery. *IEEE Access*. (2019) 7:112026–36. doi: 10.1109/ACCESS.2019.2933368

Conflict of Interest: CH is a co-inventor on patents related to physiological waveform assessment and variability analysis. MF holds patents on fetal EEG and ECG monitoring: US9,215,999 and WO2018160890.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

SYSTEMATIC REVIEW
published: 02 August 2021
doi: 10.3389/fped.2021.661400

Systematic Review of Intrapartum Fetal Heart Rate Spectral Analysis and an Application in the Detection of Fetal Acidemia

Luísa Castro [1,2,3], Maria Loureiro [4,5], Teresa S. Henriques [1,2] and Inês Nunes [1,5,6]*

[1] Faculty of Medicine, Centre for Health Technology and Services Research (CINTESIS), University of Porto, Porto, Portugal, [2] Health Information and Decision Sciences Department - MEDCIDS, Faculty of Medicine, University of Porto, Porto, Portugal, [3] School of Health of the Polytechnic of Porto, Porto, Portugal, [4] Faculty of Engineering, University of Porto, Porto, Portugal, [5] Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal, [6] Centro Materno-Infantil do Norte – Centro Hospitalar e Universitário do Porto, Porto, Portugal

It is fundamental to diagnose fetal acidemia as early as possible, allowing adequate obstetrical interventions to prevent brain damage or perinatal death. The visual analysis of cardiotocography traces has been complemented by computerized methods in order to overcome some of its limitations in the screening of fetal hypoxia/acidemia. Spectral analysis has been proposed by several studies exploring fetal heart rate recordings while referring to a great variety of frequency bands for integrating the power spectrum. In this paper, the main goal was to systematically review the spectral bands reported in intrapartum fetal heart rate studies and to evaluate their performance in detecting fetal acidemia/hypoxia. A total of 176 articles were reviewed, from MEDLINE, and 26 were included for the extraction of frequency bands and other relevant methodological information. An open-access fetal heart rate database was used, with recordings of the last half an hour of labor of 246 fetuses. Four different umbilical artery pH cutoffs were considered for fetuses' classification into acidemic or non-acidemic: 7.05, 7.10, 7.15, and 7.20. The area under the receiver operating characteristic curve (AUROC) was used to quantify the frequency bands' ability to distinguish acidemic fetuses. Bands referring to low frequencies, mainly associated with neural sympathetic activity, were the best at detecting acidemic fetuses, with the more severe definition ($pH \leq 7.05$) attaining the highest values for the AUROC. This study shows that the power spectrum analysis of the fetal heart rate is a simple and powerful tool that may become an adjunctive method to CTG, helping healthcare professionals to accurately identify fetuses at risk of intrapartum hypoxia and to implement timely obstetrical interventions to reduce the incidence of related adverse perinatal outcomes.

Keywords: fetal acidemia, fetal heart rate, spectral analysis, frequency bands, intrapartum

OPEN ACCESS

Edited by:
Anup C. Katheria, Sharp Mary Birch Hospital for Women & Newborns, United States

Reviewed by:
Philip Warrick, PeriGen Inc., Canada
Deepak Chawla, Government Medical College and Hospital, India

****Correspondence:***
Luísa Castro
luisacastro@med.up.pt

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *16 April 2021*
Accepted: *01 July 2021*
Published: *02 August 2021*

Citation:
Castro L, Loureiro M, Henriques TS and Nunes I (2021) Systematic Review of Intrapartum Fetal Heart Rate Spectral Analysis and an Application in the Detection of Fetal Acidemia. Front. Pediatr. 9:661400. doi: 10.3389/fped.2021.661400

INTRODUCTION

The fetus depends on the mother for oxygen and carbon dioxide exchange and for glucose supply, through the placenta, to maintain aerobic metabolism and adequate energy production. If there is a problem in maternal circulation, maternal respiratory function, placental blood supply, blood gas exchange within the placenta, or fetal and umbilical circulation, the fetus may suffer hypoxia. This

decreased oxygen concentration in fetal arterial blood—hypoxemia—can lead to decreased oxygen in the tissues—hypoxia/metabolic acidosis—which, if not reversed, may cause a series of long-term sequelae, such as hypoxic-ischemic encephalopathy – (HIE) leading to the later development of cerebral palsy, or perinatal death (1). In order to prevent brain damage or perinatal death, it is important to diagnose hypoxia as soon as possible, allowing obstetrical interventions to reverse it before irreparable damage develops. To make this possible and practical to screen for, intrapartum fetal monitoring with cardiotocography (CTG), a technique which evaluates fetal heart rate (FHR) and uterine contractions, has been widely used to help detect fetal hypoxia/acidosis and prevent related adverse perinatal outcomes (2).

FHR signals can be extracted using external or internal methods. The most common mode of acquisition in clinical practice is non-invasive and consists of using a Doppler probe applied on the maternal abdomen. In some situations, invasive FHR acquisition through an electrode placed on the fetal scalp (FSE), using fetal electrocardiogram, is preferred, enabling a more effective signal acquisition, when a satisfactory trace cannot be obtained by external FHR, for example, due to maternal obesity. FSE can only be used during labor, after the rupture of membranes and the beginning of cervical dilatation (3). Although not yet fully available in clinical practice, another promising technique is the transabdominal electrocardiogram which is more reliable than the ultrasound to acquire FHR signals, but still having some limitations in separating the fetal from the maternal heart rate signal. At last, fetal magnetocardiography (fMCG), the magnetic analog of fetal ECG, used in research settings and a highly effective method to evaluate fetal arrhythmias, uses the complexity of Superconducting Quantum Interference Device (SQUID) technology and is very expensive, therefore not being suitable for clinical routine use (4).

CTG visual analysis has limitations regarding its validity, reproducibility, and inter and intra-observer agreement due to the complex nature of FHR traces (5, 6). In addition, it can sometimes lead to unnecessary obstetrical interventions, such as cesarean section and operative vaginal delivery, which are associated with maternal and perinatal risks (7–9). To tackle these problems, attempts have been made to analyze CTGs with computerized systems, automating the diagnosis of fetal hypoxia/acidosis (10). Linear and non-linear algorithms have been applied to FHR signals aiming at improving performance results in the prediction of acidemia (11–13). However, although promising results have been achieved with non-statistical and complex indices (14–16), only statistical tools have been included in clinical practice, probably enabling a more intuitive interpretation by the practicing obstetrician. A crucial feature of the FHR traces is variability, which is physiologically related to the fetal heart autonomous nervous system control (17). When compared to statistical indices, spectral analysis has been referred to as providing a more appropriate evaluation of the FHR non-linear periodic changes (18). As such, spectral analysis of the FHR has been extensively explored for the detection of several pathological conditions in the fetus (19–23).

Spectral analysis determines the energy in specific frequency components of heart rate variability, and changes in the power distribution of the FHR have been considered as a predictor of fetal distress in antepartum and intrapartum settings, reflecting the autonomic nervous system (ANS) activity (24–26), which is activated by hypoxemia (25, 27). Parasympathetic and sympathetic systems compose the two divisions of the ANS, which regulates the body's unconscious actions. While the parasympathetic system acts in *rest-and-digest* and *feed-and-breed* situations, the sympathetic system acts in *fight-or-flight* situations. This method takes into consideration various frequency domains that cause complex heart rate changes and are related to diverse systems of the human body (28). Although several promising results were observed in various studies using this technique, comparison problems arise when different authors use a disparate selection of frequency bands to integrate the spectrum (24).

Several methods have been proposed for power spectral estimation of real time series, which basically fall into two groups: non-parametric techniques such as the periodogram and the Welch method (modified periodogram), and parametric techniques which identify a model representing a good approximation of the sampled signal, such as autoregressive models. Autoregressive models are defined by some order, and their coefficients can be estimated by the least-squares procedure or through the Yule-Walker equations. While all spectral methods provide only estimates of the real power spectral density (PSD) of the signal and comparable results, non-parametric methods have the advantages of simplifying the algorithm and its high processing speed. On the other hand, parametric methods provide accurate PSD estimation even on a small number of samples but require the verification of the selected model's suitability and its order or complexity (29, 30).

It is known that in the FHR power spectrum, one observes four regions, different from the components observed in the power spectrum of an adult heart rate signal (17, 29). These four components in the spectrum of the FHR have been proposed in 2003 by Signorini et al. (17) and adopted by several studies (3, 31–35). The lowest is the very low frequency (VLF), ranging from 0 to 0.03 Hz, being related to long period events and non-linear contributions; followed by the low frequency (LF), ranging from 0.03 to 0.15 Hz and being mainly associated with neural sympathetic activity. These two low frequency bands, VLF and LF, are also associated with clinical variability such as accelerations and decelerations. The next is a typical band of the FHR spectrum, the movement frequency (MF), located between 0.15 and 0.5 Hz, marking maternal breathing and fetal movements (36, 37). The high frequency (HF) ranges from 0.5 to 1 Hz and indicates the presence of fetal breathing. Other parameters are recommended besides these specific spectral power bands, such as FHR mean and variance values and the LF/(HF + MF) ratio. This ratio quantifies the autonomic

Abbreviations: AUROC, area under the receiver operating characteristic curve; CTG, cardiotocography; FHR, fetal heart rate; HF, high frequency; LF, low frequency; LLF, low low frequency; MF, movement frequency; PSD, power spectral density; ROC, receiver operating characteristic; VLF, very low frequency.

balance between neural control mechanisms from sympathetic and parasympathetic origins (17), in parallel with the ratio formulated as LF/HF, normally calculated in adults (29).

Thereby, the primary goal of this work was to determine which frequency bands are the best at screening fetal hypoxia/acidosis. First, a systematic review of the literature was conducted to collect all the different frequency domain bands used for FHR spectral analysis. Then, the power of the different bands found was computed on an open-access dataset using non-parametric spectral analysis in order to address their performance at screening fetal hypoxia/ acidosis.

MATERIALS AND METHODS

Systematic Review of FHR Spectral Analysis

A systematic review was conducted in order to obtain a collection of the original papers reporting FHR spectral analysis for revision of the spectral bands employed. All original research articles reporting spectral power analysis of real human fetal heart rate (or RR intervals) signals, published until the end of June 2020, and available in PubMed, were included in the analysis. For this purpose, papers were obtained from an online search on PubMed/Medline with the following query:

("spectral"[All Fields] OR "spectrally"[All Fields] OR "frequency domain"[All Fields] OR "frequency analysis"[All Fields] OR "frequency parameters"[All Fields] OR "linear models"[All Fields])

AND "fetal heart rate"[All Fields]) AND ("human s"[All Fields] OR "humans"[MeSH Terms] OR "humans"[All Fields] OR "human"[All Fields]).

From the set of papers obtained with the query described, two researchers, familiar with signal analysis, selected the papers to include. This selection was based on the title and abstract only, and performed independently, by each researcher. The papers to include were then distributed randomly to three of the authors who extracted, from the full paper, the relevant information for this study. During full revision of the studies, referenced papers, which were not captured by the query, were also manually selected. Previous studies have shown that the distribution of the frequency spectrum suffers changes with the baby maturation (38), hence the revision was limited to articles reporting analysis of signals referring to the term of pregnancy and to the intrapartum period. The exclusion criteria were: language other than English; no report of signal analysis (mainly reviews); reported analysis in simulated signals or from non-human fetuses; reported analysis on signals obtained from magnetocardiography (by its lack of applicability in clinical practice); and used data from the antepartum period. In the end, a consensus was reached about the papers that accurately reported spectral analysis of intrapartum FHR and that were included in the final revision. The information extracted from the studies, beyond the frequency bands used, included, if available, the employed spectral analysis method, the interpretation of the bands used, the type of signal spectrally analyzed (FHR or RR intervals) and its sampling frequency, and the signal acquisition mode.

A total of 172 abstracts were identified for screening, selected automatically by the query, and 3 additional articles were obtained from reference checking during the eligibility step. After exclusions, 26 articles remained to be included in the review (**Figure 1**). This systematic review focused on the frequency spectral bands being used and reported for spectral analysis of FHR in the literature. It was not in the scope of this research to consider other information from those articles, commonly addressed in systematic reviews such as main outcomes, discriminatory validity or other results. In this research, the information retrieved from the reviewed studies concerned only the methodology regarding the spectral bands used.

Application in Fetal Acidemia Detection

FHR Data

In order to compare the performance of the various spectral indices used in the articles reviewed, the first open-access database for research analysis in intrapartum CTG was used, enabling the objective comparison between results from different methods (39). The open-access database CTU-UHB consists of a total of 552 intrapartum FHR recordings obtained, between April 2010 and August 2012, at the labor ward of the University Hospital in Brno, Czech Republic. Along with the intrapartum CTGs, relevant fetal (including biochemical parameters of umbilical arterial blood samples) and maternal clinical data were available (39). All recordings refer to singleton pregnancies, with more than 36 completed weeks. The CTU-UHB recordings start no more than 90 min before the delivery, and each is at most 90 min long, sampled at 4 Hz (39).

Before labor, the arterial pH of a healthy fetus is around 7.35, whereas, at birth, the average pH of the umbilical artery blood is around 7.25. In this sense, moderate neonatal acidosis will occur when the pH is at least below 7.15 (40). However, not all fetuses in this situation are at risk for immediate or long-term complications, which will depend on the type of acidosis and its severity. Respiratory acidosis *per se* does not carry long-term neurological complications, while metabolic acidosis is related to prolonged hypoxia and needs more time to reverse, even after hypoxia is corrected. By its turn, prolonged metabolic acidosis, involving lower pH values, is associated with irreversible organ damage (40). Fetal metabolic acidosis is defined by an umbilical artery pH below 7.00 and a base-deficit in the extracellular fluid (BD_{ecf}) above 12 mmol/l (24, 40, 41). However, there is already an association with adverse short-term perinatal outcomes when pH values are below 7.05 and the BD_{ecf} is above 10 mmol/l (40). Lower arterial pH values are not common in regular cross-sectional studies or even in large randomized controlled trials conducted in the intrapartum setting and considering fetal acidemia as a main outcome measure (42, 43), being a rare event in nowadays' clinical practice (incidence lower than 0.6%) (42), in developed countries. Thus, cutoffs such as 7.20 (31), 7.15 (28), 7.10 (31), and 7.05 (35, 44) are regularly employed in studies of fetal acidemia/hypoxia detection. Accordingly, in this study, four different arterial pH cutoffs were considered for the definition of fetal acidemia: 7.05, 7.10, 7.15, and 7.20. The 7.00 pH cutoff was

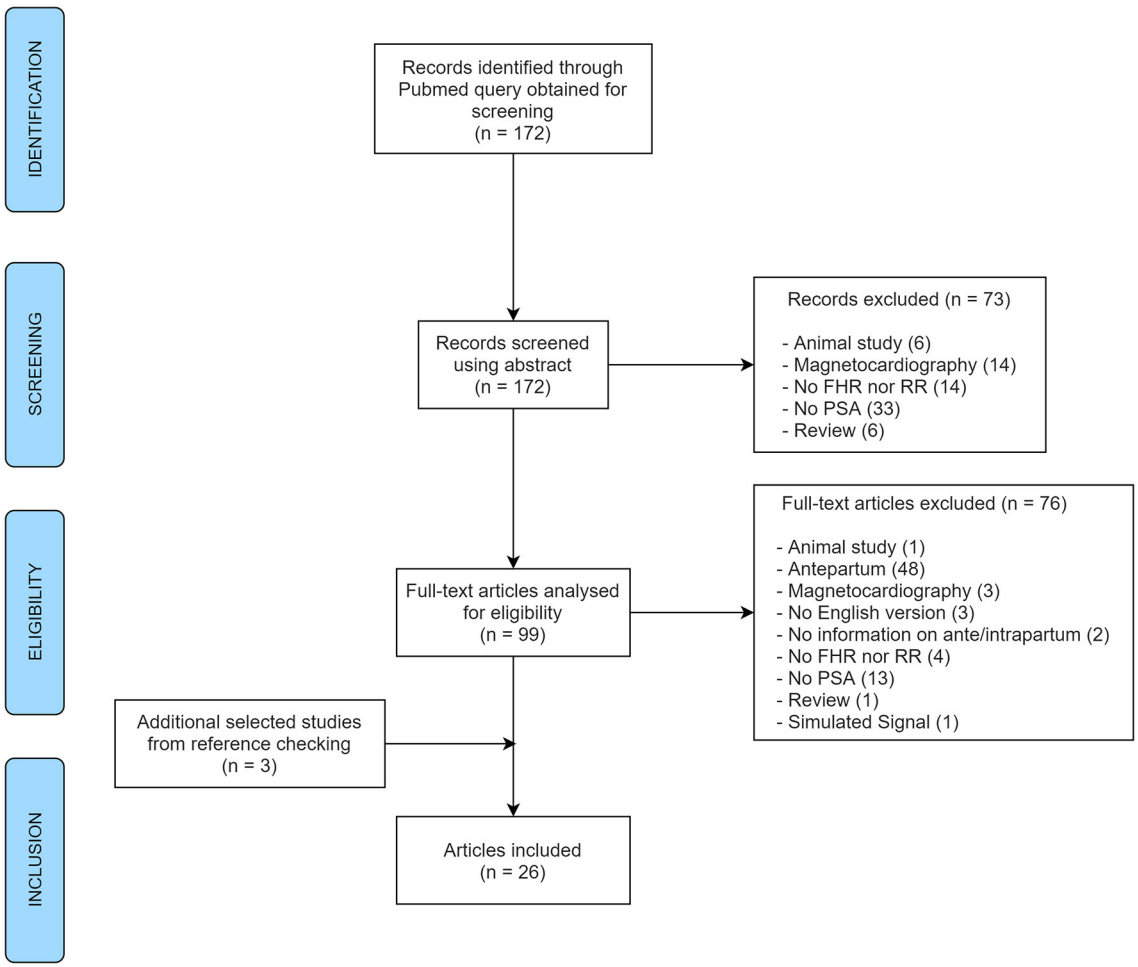

FIGURE 1 | PRISMA flow diagram of the article selection process for the systematic review (FHR, fetal heart rate; RR, interval between R peaks; PSA, power spectral analysis).

not considered for this analysis since only three of the included CTGs referred to newborns with an arterial pH below 7.00 (incidence of 0.54%, in line with the aforementioned literature).

Signal Pre-processing

Pre-processing of the FHR signal was required to remove artifacts and signal loss that is commonly present in the final minutes before delivery, related to the movements of the descending fetus and the pregnant woman who is pushing and the improper placement of the sensors. The algorithm used, described in detail in (3), detects samples lower than 60 bpm and above 200 bpm and consecutive differences higher than 25 bpm. The segments detected were then replaced by linear interpolation if referring to <2 s. For longer periods, segments were substituted by the previous segment of the same length. All samples were rounded to units. After discarding the last 5 min of each recording, the final 30 min were selected for the following analyses. After pre-processing the available signal, FHR recordings with signal loss above 15% in the last hour (excluding the last 5 min) were rejected, ending with 246 traces selected for spectral analysis.

Spectral Analysis

Parseval's theorem states that the total energy of a signal in the time domain is equal to its energy in the frequency domain and, for this reason, the square of the magnitude of the Fourier Transform of a signal represents its power density (22). Using the Fast Fourier Transform (FFT) and the frequency domain bands reported in the literature, various indices of FHR's power spectral density (PSD) were computed.

To estimate the PSD of the FHR, a non-parametric method was employed, the Welch method, with a window length of 256 samples with a 62.5% overlap, similarly to previous studies (3, 11, 13, 31, 34). The Welch periodogram-based estimation method corresponds to integrating the periodogram (using the rectangular method) in windows of overlapped segments, modified periodograms, that are averaged to obtain the PSD estimate, allowing an estimation of the average power of the signal. The integration was done between limited frequency ranges. The FHR power percentages were computed, following previous recommendations (45), corresponding to the power of each frequency band divided by the average power of the

whole signal (46). These normalized values allow the detection of relative changes, instead of changes in the total power, that would mask the first ones (25).

Computing a PSD estimation directly on the FHR signal results in a high impulse around 0 Hz frequency, the DC component, corresponding to the signal's average value over a period, disguising relevant signals with relatively small amplitude. Thus, after preprocessing the signals, and before the spectrum estimation, the mean FHR was removed for each signal to reduce the DC-offset of FHR. All preprocessing computations and spectral analysis were done using MATLAB (R2020b, MathWorks, Natick, MA, USA).

Statistical Analysis

Descriptive measures and statistical tests were employed to properly compare the FHR power percentage of the acidemic and the non-acidemic fetuses on various frequency bands. The normality of quantitative variables was verified visually and confirmed with the Kolmogorov-Smirnov test. Since almost all indices could not be assumed as normally distributed, median and interquartile interval (first quartile, Q1 - third quartile, Q3) were employed. Categorical variables were described by absolute and relative frequencies.

For the comparison between acidemic and non-acidemic groups of fetuses, the Mann-Whitney test was employed, comparing the distribution functions of the spectral indices between the two groups. Subsequently, and for the frequency bands where significant differences were found, a receiver operating characteristic (ROC) curve was computed (47). The spectral indices' ability to distinguish acidemic fetuses from non-acidemic ones was evaluated using the area under the ROC curve (AUROC). These curves relate sensitivity—ability to recognize acidemic fetuses correctly—and specificity—ability to identify non-acidemic fetuses—for each cutoff of the discriminating index (31), or power spectral band, in our case.

For descriptive and inference statistics, SPSS Statistics (v.25; IBM SPSS, Chicago, IL) was used. For all statistical tests, a significance level of 0.05 was pre-defined.

RESULTS

Literature Review of Spectral Analysis of FHR

Overall, a total of 176 papers were reviewed by abstract. At the abstract screening step, 73 papers were excluded for several reasons: 6 referring to animal studies, 6 to reviews (with no signal analyses), 14 did not analyze FHR nor RR, 33 did not apply power spectral analysis at all, and finally, 14 were excluded as they used magnetocardiography for FHR acquisition. As for the analyses of full-text articles for eligibility, another 76 documents were excluded for the same set of reasons described in the previous step (**Figure 1**) together with: one article which referred to the analysis of simulated signal and 48 studies reporting analysis of antepartum signals (plus two with no information). Three additional articles obtained from reference checking were also included. In the end, 26 different original papers reporting spectral analysis of intrapartum FHR (or RR intervals) were included in the analysis, from 1975 to 2019.

The most frequent aim in the studies included was the assessment of acidemia state alone ($n = 14$) or combined with other characteristics of the fetus such as gender (32) or prematurity (48). There was also a study addressing the relation between PSD indices and cord arterial base deficit values at birth (49). Other than acidemia, the study of the impact of signal preprocessing for artifact correction on PSD features (50), and the effect of different acquisition modes (3), were the aim of other studies. Power spectral analysis was also employed for the characterization/detection of a great diversity of other situations: maturity (19, 20), weight or/and gender (21, 22), and conditions of fetal distress (23). Included articles described analysis on FHR and RR-intervals, mainly acquired through fetal scalp ECG ($n = 14$) and fetal Doppler transducer (external ultrasound, $n = 12$). The majority of articles ($n = 14$) reported the use of the Fast Fourier Transform method for computing the spectral density of the signals, without further details. From the remaining studies, six employed the Welch method, three the autoregressive model, one used the matching pursuit technique (51), and the other used the standard periodogram method (35). Regarding signal and sample rate, the great majority of the articles included analyzed FHR ($n = 22$) signals sampled at 4 Hz ($n = 10$) or 2 Hz ($n = 5$). Other sampling rates were also analyzed, although less frequently, such as 1, 10, and 16 Hz. From the 4 studies reporting power spectral analysis on RR signals, three used signals sampled at 4 Hz.

To address the main aim of this study, all the frequency bands used in the power spectral analysis were extracted from the 26 articles included in the revision. A wide variety of frequency bands' limits was found, and only two articles did not report the use of any specific interval or frequency bands: one from 1975 aiming at characterizing spectral density for premature infants or prolonged labor (19) and a more recent one aiming at computing FHR from beat-to-beat signal (52). From the remaining 24 (**Table 1**), regarding very low frequency, the bands' limits employed were 0–0.03 Hz in seven articles, 0–0.04 Hz in three, and 0.003–0.004 Hz in one study. In contrast, thirteen studies did not refer to the very low frequency power energy. From the eleven studies reporting the power in the very low frequency band, six attributed its activity to the thermoregulation and slow regulating systems of peripheral vessels, as stated in 1996 Task Force (29). One study reported the use of an intermediate frequency, named as low low frequency (LLF), referring to the range 0.04–0.08 Hz (51).

The low frequency was addressed in all the 24 articles, where the most common band was within the interval 0.04–0.15 Hz, used in 12 studies, followed by 0.03–0.15 Hz, addressed by 8 articles. Regarding the power in the LF band, most articles attribute it to the activity of the sympathetic system (3, 28, 31, 32, 34, 49, 59, 60, 62), while others link it to the combined activity of sympathetic and parasympathetic activity (11, 13, 20, 21, 44, 48, 57, 61). This different interpretation cannot, in some cases at least, be attributed to the different frequency limits of the bands used for the LF computation (check **Table 1**), and is

TABLE 1 | Information on different frequency bands and related contents of the 24 articles reviewed.

Frequency band (Hz)	Signal-sampling rate	PSD method	Main goal	Origin of the bands (and their references)	Publication year (article reference)
VLF (0–0.04) LF (0.04–0.15) HF (0.15–0.4) VHF (0.75–1.5)	FHR – 2 Hz	FFT	Diagnose fetal acidemia	Adults	2001 (28)
LF (0.03–0.07) MF (0.07–0.13) HF (0.13–1)	FHR – 4 Hz	FFT	Association PSD with fetal cord arterial base deficit values at birth	Adults (53–55)	2001 (49)
LF (0.04–0.15) HF (0.15–1.0)	FHR – 16 Hz	FFT	Diagnose fetal acidemia	Adults	2005 (44)
	RR – 500 Hz				2007 (56)
	FHR – 16Hz				2013 (57)
LF (0.04–0.15) HF (0.15–0.4)	FHR – 10 Hz	Welch method	Diagnose fetal acidemia	Adults	2015 (11)
VLF (0.003–0.04) LF (0.04–0.15) HF (0.15–0.4)	FHR – 10 Hz	Welch method	Diagnose fetal acidemia	Adults	2017 (13)
VLF (0–0.04) LF (0.04–0.15) HF_A (0.15–0.4) HF_N (0.4–1.5)	RR – 4 Hz	FFT	Compare real signal with artifacts and without artifacts	Neonates and Adults	2008 (50)
LF (0.04–0.15) HF (0.4–1.5)	FHR – 4 Hz	FFT	Compare near to post term during active and quiet sleep	Neonates (45, 58)	2009 (20)
	RR – 4 Hz	FFT	Diagnose fetal acidemia		2010 (27)
		FFT			2011 (59)
VLF (0–0.03) LF (0.03–0.15) MF (0.15–0.5) HF (0.5–1)	FHR – 2 Hz	Welch method	Diagnose fetal acidemia	Fetuses (Antepartum) (17)	2006 (31)
	FHR – 2 and 4 Hz	Welch method	Internal vs. external acquisition modes		2006 (3)
	FHR – 4 Hz	Not reported	Diagnose Fetal Acidemia/gender differences		2009 (32)
		Autoregressive model	Diagnose fetal acidemia		2011 (33)
		Welch method	Indices in beat-to-beat vs. 4 Hz		2013 (34)
		Welch method	Differences in fetal gender		2017 (60)
		Periodogram	Diagnose fetal acidemia		2019 (35)
LF (0.03–0.15) MF (0.15–0.5) HF (0.5–1)	FHR – 4 Hz	Autoregressive model	Diagnose fetal acidemia	Fetuses (Antepartum) (17)	2010 (61)
LF (0.04–0.15) MF (0.15–0.5) HF (0.5–1)	FHR – 2 Hz	FFT	Diagnose fetal acidemia and prematurity	Fetuses (Antepartum) (17)	2012 (48)
			Differences in fetal weight and gender		2014 (21)
VLF (0–0.04) LLF (0.04–0.08) LF (0.08–0.15) HF (>0.15)	FHR – 1 Hz	Matching pursuit	Diagnose fetal acidemia	Not mentioned	2006 (51)
LF (0.03125–0.1)	FHR – 2 Hz	FFT	Diagnose fetal distress	Not mentioned	2010 (23)
LF (0.02–0.14) MF (0.1–0.4) HF (0.4–1.4)	FHR – 4 Hz	Autoregressive model	Diagnose fetal acidemia	Not mentioned	2015 (62)

VLF, very low frequency; LLF, low low frequency; LF, low frequency; MF, movement frequency; HF, high frequency; VHF, very high frequency; FHR, fetal heart rate; RR, interval between R peaks; PSD, power spectral density; FFT, fast Forrier transform.

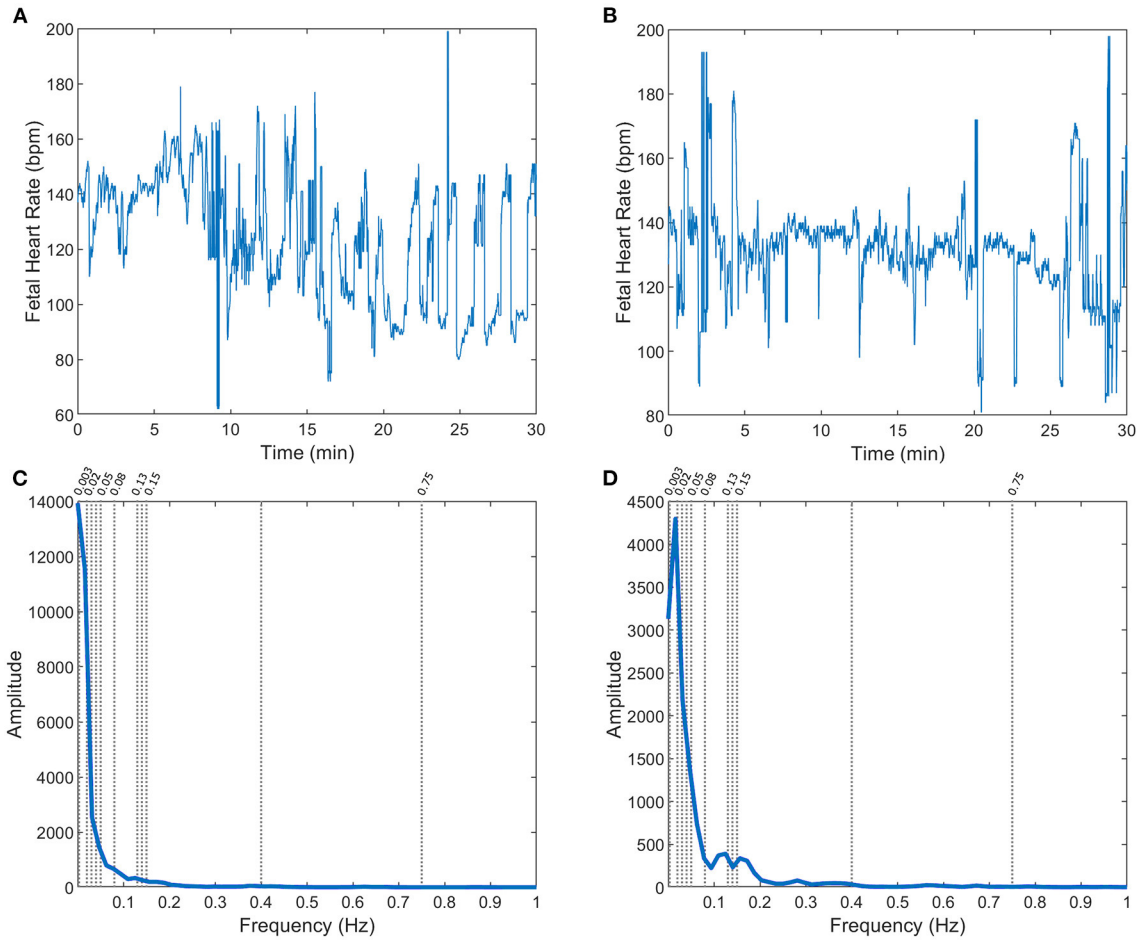

FIGURE 2 | FHR **(A,B)** and respective power spectra (**C,D**, with specific frequency limits) for two fetuses in the CTG database: with arterial blood pH = 7.02 **(A,C)** and pH = 7.2 **(B,D)**, at birth.

probably inherited by the same controversial interpretation of the LF component in adults' heart rate analysis (29).

Only 12 studies assessed the power in the movement frequency, and ten of them computed its energy within the limits 0.15–0.5 Hz. One article reported limits of 0.07–0.13 Hz for the MF band (49). Fetal movements and maternal breathing are the activities most associated to the MF band (3, 21, 34, 62), with only one article (49) attributing both sympathetic and parasympathetic nervous control to the activity in this frequency band. This difference in the interpretation can be attributed to the different ranges in the MF band, which shows a large superposition with the LF component.

The high frequency band was employed in 23 studies, and the interval 0.5–1 Hz was the most used, being reported in 10 articles. Only one study mentioned the use of a very high frequency band, within 0.75–1.5 Hz, aiming to the detection of fetal acidemia (28). The majority of the articles addressing the HF band interpreted its power as reflecting the activity on the parasympathetic systems, as advocated by the 1996 Task Force (29), and only a few specify the connection to fetal breathing (61, 62) due to the vagal nerve (28). The VHF was reported once, interpreted as reflecting the domain of the heart beating (28).

Regarding the ratio of spectral bands, interpreted as the sympathovagal balance in the control of heart rate activity (between the parasympathetic and the sympathetic branches), three of the included articles reported the use of LF/(MF + HF) and eight, who did not report the use of movement frequency band [with one exception (21)], preferred LF/HF. Three studies reported both ratio measures LF/(MF + HF) and LF/HF (32, 34, 48), referring that the two ratios provide different balances between the autonomic nervous system branches.

No clear relation was found between the type of signal analyzed or sampling rate regarding the bands adopted.

From the 24 papers using specific frequency bands, seven refer to bands inspired in human adults' signals, four in newborns and ten used bands defined based on power spectral analysis of the FHR, spanning a wide range of publication years (2001–2017 for adult bands and 2006–2019 for fetal analysis inspired bands) (**Table 1**).

Spectral Analysis on CTG Intrapartum Open-Access Database

A total of 246 FHR recordings were included in the analysis, referring to the last half an hour (excluding the last 5 min) of the recordings from the CTU-UHB open-access database, with <15% of signal loss. The FHR signals were slightly more often from female fetuses (51%), with occipital presentation (91%), a median (Q1–Q3) of 40 (39–41) weeks of gestation, and mother's age of 29 (26.8–32) years old. Only 7, 12, 39, and 72 cases were classified as acidemic babies, considering arterial blood pH cutoffs of 7.05, 7.10, 7.15, and 7.20, respectively (**Figure 2** - two FHR examples with the respective power spectral estimate).

Descriptive statistics of the several spectral indices computed on the 246 CTU-UHB recordings are presented in **Supplementary Tables 1–4**, together with their comparison between acidemic and non-acidemic groups of fetuses. Acidemic groups were defined for each pH cutoff, namely pH $\leq$ 7.20, pH $\leq$ 7.15, pH $\leq$ 7.10, and pH $\leq$ 7.05. Overall, acidemic cases pointed to higher values of VLF power and lower powers in all other spectral bands (LF, LLF, MF, HF, and VHF ranges), and significant differences were found between the two groups defined by all four pH cutoffs, for some spectral bands (**Supplementary Tables 1–3**). In particular, for cutoff 7.20, significant differences were found only for VHF band 0.75–1.5 Hz (28) and a low AUROC of 0.593 was obtained (see **Table 2**). For cutoff 7.15, significant differences between acidemic and non-acidemic groups were detected for VLF, LLF, LF, and VHF indices (**Supplementary Table 2**), and their AUROC were computed (**Table 3**). For those, the best AUROC values were: 0.626 for LF 0.04–0.15 Hz (11, 13, 20, 21, 27, 28, 44, 48, 50, 56, 57, 59) and 0.624 for LLF 0.04–0.08 Hz (51) and for LF 0.03125–0.1 Hz (23), as depicted in **Table 3**.

For cutoff 7.10, significant differences were found for all spectral bands' indices (**Supplementary Table 3**), and highest AUROC values were achieved, comparing with the previous acidemic definition (see **Table 4**). Overall, the best AUROC values were: 0.731 for LF 0.04–0.15 Hz (11, 13, 20, 21, 27, 28, 44, 48, 50, 56, 57, 59); 0.730 for LF 0.03125–0.1 Hz (23) and 0.729 for LLF 0.04–0.08 Hz (51) (see **Table 4**).

For cutoff 7.05, although the acidemic group had only 7 subjects, significant differences were found for VLF and LF spectral bands (**Supplementary Table 4**), and their AUROC were computed (**Table 5**). Overall, the best AUROC values were: 0.770 for LF 0.03–0.07 Hz (49); 0.763 for LF 0.02–0.14 Hz (62) and 0.762 for LF 0.03–0.15 Hz (3, 31–35, 60) (see **Table 5**).

TABLE 2 | Area under ROC curve and correspondent non-parametric confidence interval, for the bands significantly different between groups of fetal acidemia defined with pH cutoff of 7.20, in the CTU-UHB open-access database (section Spectral Analysis on CTG Intrapartum Open-Access Database).

Frequency band (Hz) (reference)	Area	95% confidence interval
VHF (0.75–1.5) (28)	0.593	(0.514–0.672)

VHF, very high frequency.

DISCUSSION

A total of 26 articles were included in the systematic review presented here, surveying for spectral analysis of intrapartum fetal heart rate. As suspected, a great panoply of frequency bands has been applied, some inspired in fetal heart rate spectrum evidence and others in adult and neonatal heart rate studies. Although it seems, from our results, that most recent studies preferably select frequency bands inspired in the fetus, as normally expected, we also found some recent studies choosing a spectrum of adult-derived bands. This finding might reflect that there is some controversy regarding the proper bands to use in fetuses (11, 13). In a study comparing the performance of spectral analysis and the Hurst parameter for fetal acidemia detection, the Hurst parameter revealed to be a potential marker of fetal acidosis, overcoming the performance of the spectral index LF/HF, with the advantage of not depending on the choice of the partitioning of frequency bands (11). This interest in providing an alternative to the spectrum splitting is also fostered by the fact that several conditions impact the association between the power in the spectral bands and ANS activity in the fetus, such as maturity differences between the sympathetic and parasympathetic systems; uterine contractions affecting the intrathoracic fetal pressure, which influence the FHR absolute values and variability (63, 64). As noticed in a 2008 review by Van Laar et al. (45), the different heart rate and pattern of breathing movements of fetuses compared to adults suggest that adult-inspired frequency bands may not be perfectly chosen for FHR analysis, and their recommendation of an agreement on the frequency bands

TABLE 3 | Areas under ROC curves and correspondent non-parametric confidence interval, for the band significantly different between groups of fetal acidemia defined with pH cutoff of 7.15, in the CTU-UHB open-access database (section Spectral Analysis on CTG Intrapartum Open-Access Database).

Frequency band (Hz) (references)	Area	95% confidence interval
VLF (0–0.03) (3, 31–35, 60)	0.610	(0.521–0.699)
LLF (0.04–0.08) (51)	**0.624**	**(0.533–0.716)**
LF (0.02–0.14) (62)	0.617	(0.520–0.714)
LF (0.03–0.07) (49)	0.617	(0.521–0.713)
LF (0.03–0.15) (3, 31–35, 60, 61)	0.617	(0.521–0.714)
LF (0.03125–0.1) (23)	**0.624**	**(0.532–0.715)**
LF (0.04–0.15) (11, 13, 20, 21, 26–28, 44, 48, 50, 56, 57, 59)	**0.626**	**(0.534–0.717)**
VHF (0.75–1.5) (28)	0.615	(0.519–0.710)

In bold are the three best AUROC values. VLF, very low frequency; LLF, low low frequency; LF, low frequency; VHF, very high frequency.

TABLE 4 | Areas under ROC curves and correspondent non-parametric confidence interval, for the bands significantly different between groups of fetal acidemia defined with pH cutoff of 7.10, in the CTU-UHB open-access database (section Spectral Analysis on CTG Intrapartum Open-Access Database).

Frequency band (Hz) (references)	Area	95% confidence interval
VLF (0–0.03) (3, 31–35, 60)	0.724	(0.597–0.852)
VLF (0–0.04) (28, 50, 51)	0.717	(0.584–0.851)
VLF (0.003–0.04) (13)	0.717	(0.584–0.851)
LLF (0.04–0.08) (51)	**0.729**	**(0.598–0.860)**
LF (0.02-0.14) (62)	0.703	(0.544–0.863)
LF (0.03–0.07) (49)	0.700	(0.542–0.859)
LF (0.03–0.15) (3, 31–35, 60, 61)	0.703	(0.543–0.864)
LF (0.03125-0.1) (23)	**0.730**	**(0.601–0.859)**
LF (0.04–0.15) (11, 13, 20, 21, 26–28, 44, 48, 50, 56, 57, 59)	**0.731**	**(0.605–0.856)**
LF (0.08−0.15) (51)	0.710	(0.583–0.837)
MF (0.07−0.13) (49)	0.722	(0.597–0.846)
MF (0.1-0.4) (62)	0.698	(0.564–0.832)
MF (0.15–0.5) (3, 21, 31–35, 48, 60, 61)	0.680	(0.542–0.818)
HF (>0.15) (51)	0.684	(0.538–0.829)
HF (0.13-1) (49)	0.677	(0.533–0.821)
HF (0.15–0.4) (11, 13, 28)	0.683	(0.548–0.818)
HF (0.15–1.0) (44, 56, 57)	0.679	(0.535–0.824)
HF (0.4–1.5) (20, 27, 50, 59)	0.678	(0.512–0.844)
HF (0.4–1.4) (62)	0.675	(0.509–0.841)
HF (0.5–1) (3, 21, 31–35, 48, 61)	0.673	(0.511–0.835)
VHF (0.75–1.5) (28)	0.702	(0.540–0.865)

In bold are the three best AUROC values. VLF, very low frequency; LLF, low low frequency; LF, low frequency; MF, movement frequency; HF, high frequency; VHF, very high frequency.

chosen remains present. Nevertheless, several studies have been conducted for fetal acidosis detection, wherefrom all the bands and ratios used, the VLF and LF bands are revealed as promising in separating acidemic groups (13, 27, 33, 35, 51, 59).

TABLE 5 | Areas under ROC curves and correspondent non-parametric confidence interval, for the bands significantly different between groups of fetal acidemia defined with pH cutoff of 7.05, in the CTU-UHB open-access database (section Spectral Analysis on CTG Intrapartum Open-Access Database).

Frequency band (Hz) (references)	Area	95% confidence interval
VLF (0–0.03) (3, 31–35)	0.692	(0.514–0.870)
LLF (0.04–0.08) (51)	0.759	(0.617–0.900)
LF (0.02-0.14) (62)	**0.763**	**(0.593–0.932)**
LF (0.03–0.07) (49)	**0.770**	**(0.608–0.932)**
LF (0.03–0.15) (3, 31–35, 60, 61)	**0.762**	**(0.589–0.936)**
LF (0.03125-0.1) (23)	0.759	(0.616–0.902)
LF (0.04–0.15) (11, 13, 20, 21, 26–28, 44, 48, 50, 56, 57, 59)	0.759	(0.611–0.906)

In bold are the three best AUROC values. VLF, very low frequency; LLF, low low frequency; LF, low frequency.

The early detection of acidemic fetuses was the most common aim (16/26) in the papers included in the systematic review. The capacity of each individual spectral band found in the literature, computed on the 246 FHR's signals for discriminating acidemic from non-acidemic fetuses was accessed, for four pH cutoffs: 7.05; 7.10; 7.15 and 7.20. The highest AUC values were obtained when using lower pH cutoffs. Lower cutoffs are indicative of more severe acidemia, which probably causes a more notorious change in the variability of the fetal heart rate. This allows a more efficient detection of acidemic fetuses, in particular using power spectral analysis, although it is commonly hampered by the small number of severely acidemic fetuses in available datasets (42, 43). Considering the detection of acidemic fetuses with pH $\leq$ 7.05, the best frequency bands, with AUC above 0.760, were all low frequency bands. This indicates that the low frequency band might be the best at distinguishing acidemic from non-acidemic fetuses, at least in more severe cases.

The LF bands with the highest AUC, for a cutoff of 7.05, have their origins in different studies, from fetuses, 0.03–0.15 Hz (17), 0.02–0.14 Hz (62), to adults-based, 0.03–0.07 Hz (49). Differences between these frequency ranges are small on the lower limits (0.02–0.03) but include, in the upper limits, larger differences (from 0.07 to 0.15) (17, 49). Results for this severe group should be considered with caution, due to the fact that this was a very limited set of cases, with only seven fetuses. The main limitation of this study relates to the very low incidence of severe acidemic fetuses in this FHR database, which is in accordance with previously published studies (42, 43). This limitation, combined with many band cutoffs used, has resulted in a very wide CI of the AUC. This methodology needs to be validated in a larger dataset (combination of clinical datasets) with an increased prevalence of acidemic fetuses to overcome this limitation.

For the less severe cutoffs, 7.10 and 7.15, results were similar, with the same three low frequency bands as the most promising: 0.04–0.08 (51); 0.03125–0.1 (23) and 0.04–0.15 (11, 13, 20, 21, 26–28, 44, 48, 50, 56, 57, 59). For these, the last was inspired in evidence stemming from neonatal/adult spectra (29, 45, 58), while the two previous were probably derived from empirical studies. These results do not confirm our intuition that the bands derived from FHR spectra should be more accurate for the detection of pathological cases. Indeed, the human fetus has a different heart rate and distinct pattern of breathing movements compared to a human adult or even newborns (45). Additionally, in all discriminative LF bands presented, the reduced power found in the case of acidemic cases was in accordance with previous studies (45).

Several studies (18/24) do not evaluate the VLF bands. This approach might be related to the attempt to quantify changes in the FHR baseline. To our views, clinically, in the case of sub-acute fetal hypoxia (slowly evolving hypoxia, e.g., as a result of non-reversed excessive uterine contractions during labor), the change of variability in the baseline, especially after an increased baseline shift, is very informative and therefore is not essential to evaluate LF or VLF bands. On the other hand, in acute events, such as a cord prolapse, a decrease of variability within the deceleration is specifically related to fetal hypoxia (2), and therefore the evaluation of the LF or VLF bands must be considered.

In this review, we have chosen to restrict our analysis to the intrapartum period, in an attempt to provide homogeneity to the included studies, reducing the variation of the encountered frequency bands. There is evidence that the spectral density of spectral bands is changed not only by the resting state of the fetus (38), but also, and easier to control, by the gestational age or developmental maturity (64).

Despite the multitude of definitions encountered for frequency bands, methods used and signals employed in the conducted literature review, the consistency of the results across three of the four definitions of fetal acidemia here evaluated and the performance metrics' values obtained, spectral analysis remains a powerful method to understand the dynamics of the fetal autonomic system. Studies encompassing more detailed characterization of the included subjects, namely the fetus' physiological state, should be fostered in order to develop a better performance of spectral based indices and to get them ready to be incorporated in computerized systems that aid in the clinical assessment of fetal well-being. Regarding the power spectrum of antepartum signals, and given the difficulty in providing a standard recommendation for the selection of spectral bands, the idea of partitioning the power spectrum in empirical bands, or in consecutive bands of equal bandwidth, might be a secure approach (52, 64, 65) in comparison to the compromise of choosing predefined band definitions.

Behind the efforts reflected in many studies to obtain an effective computational tool to disclose fetal distress, is the hypothesis that the frequency decomposition of the FHR can help clinicians to predict it, specifically metabolic acidosis. This could be an adjunctive methodology to the CTG that would help clinicians to implement adequate obstetrical interventions, before metabolic acidosis is severe enough to produce irreversible damages.

CONCLUSION

In conclusion, this work shows that, in general, most research papers do not agree on the definition of frequency bands to be used, which is essential for their application in clinical practice. The low frequency bands are the most promising in detecting fetuses at risk of acidosis. Furthermore, the power spectrum analysis of FHR is a powerful tool to help physicians diagnose fetal acidemia in the intrapartum period. Thus, it is essential to determine an exact diagnostic value of spectral analysis to be considered for the identification of fetuses at risk of severe acidosis and to standardize the methods used in signal pre-processing and in the spectral analysis itself to allow a better comparison between studies. In addition, to enable adequate clinical applicability, a focus on real-time intrapartum fetal monitoring should also be a priority in future research.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

LC, TH, and IN: conceptualization and methodology. ML, LC, and TH: formal analysis and investigation. ML and LC: writing—original draft preparation. ML, LC, TH, and IN: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This article was supported by National Funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., within CINTESIS, R&D Unit (reference UIDB/4255/2020).

ACKNOWLEDGMENTS

The authors acknowledge the SisPorto Research Group (Department of Obstetrics and Gynecology, Faculty of Medicine, University of Porto, Porto, Portugal) project for the scientific support, in particular we address a special thanks to Hernâni Gonçalves for many fruitful discussions on several issues related to this study, and his pertinent suggestions to improve this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2021.661400/full#supplementary-material

REFERENCES

1. Bobrow CS, Soothill PW. Causes and consequences of fetal acidosis. *Arch Dis Child Fetal Neonatal Ed.* (1999) 80:F246–F9. doi: 10.1136/fn.80.3.F246
2. Ayres-de-Campos D, Spong CY, Chandraharan E, Panel FIFMEC. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynecol Obstetr.* (2015) 131:13–24. doi: 10.1016/j.ijgo.2015.06.020
3. Gonçalves H, Rocha AP, Ayres-de-Campos DJB. Internal versus external intrapartum foetal heart rate monitoring: the effect on linear and nonlinear parameters. *Physiol Meas*. (2006)27:307–19. doi: 10.1088/0967-3334/27/3/008
4. Hoyer D, Zebrowski J, Cysarz D, Gonçalves H, Pytlik A, Amorim-Costa C, et al. Monitoring fetal maturation-objectives, techniques and indices of autonomic function. *Physiol Meas.* (2017) 38:R61–88. doi: 10.1088/1361-6579/aa5fca
5. Bernardes J, Ayres-de-Campos D. The persistent challenge of foetal heart rate monitoring. *Curr Opin Obstetr Gynecol.* (2010) 22:104–9. doi: 10.1097/GCO.0b013e328337233c
6. Ayres-de-Campos D, Bernardes J, Professor A. Inconsistencies in classification by experts and subsequent clinical decision. *BJOG Int J Obstetr Gynaecol.* (2005) 106:1307–10. doi: 10.1111/j.1471-0528.1999.tb08187.x
7. Silver RM. Implications of the first cesarean: perinatal and future reproductive health and subsequent cesareans, placentation issues, uterine rupture risk, morbidity, and mortality. *Semin Perinatol.* (2012) 36:315–23. doi: 10.1053/j.semperi.2012.04.013
8. Signore C, Klebanoff M. Neonatal morbidity and mortality after elective cesarean delivery. *Clin Perinatol.* (2008) 35:361–vi. doi: 10.1016/j.clp.2008.03.009
9. Wilmink F, Hukkelhoven C, Lunshof S, Mol BW, Van der Post J, Papatsonis D. Neonatal outcome following elective cesarean section beyond 37 weeks of gestation: a 7-year retrospective study of a national registry. *Am J Obstetr Gynecol.* (2010) 202:250.e1–8. doi: 10.1016/j.ajog.2010.01.052
10. Visser GH, Ayres-de-Campos D, Panel FIFMEC. FIGO consensus guidelines on intrapartum fetal monitoring: adjunctive technologies. *Int J Gynaecol Obstet.* (2015) 131:25–9. doi: 10.1016/j.ijgo.2015.06.021
11. 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:e0136661. doi: 10.1371/journal.pone.0136661
12. Magenes G, Bellazzi R, Malovini A, Signorini MG. Comparison of data mining techniques applied to fetal heart rate parameters for the early identification of IUGR fetuses. *Annu Int Conf IEEE Eng Med Biol Soc.* (2016) 2016:916–9. doi: 10.1109/EMBC.2016.7590850
13. 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) 21:664–71. doi: 10.1109/JBHI.2016.2546312
14. Yum M, Park EY, Kim CR, Hwang J. Alterations in irregular and fractal heart rate behavior in growth restricted fetuses. *Eur J Obstetr Gynecol Reprod Biol.* (2001) 94:51–8. doi: 10.1016/S0301-2115(00)00314-6
15. Richman J, Moorman J. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol.* (2000) 278:H2039–49. doi: 10.1152/ajpheart.2000.278.6.H2039
16. Salamalekis E, Thomopoulos P, Giannaris D, Salloum I, Vasios G, Prentza A, et al. Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilising wavelet analysis and neural networks. *BJOG Int J Obstetr Gynaecol.* (2002) 109:1137–42. doi: 10.1111/j.1471-0528.2002.01388.x
17. Signorini MG, Magenes G, Cerutti S, Arduini D. Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings. *IEEE Trans Biomed Eng.* (2003) 50:365–74. doi: 10.1109/TBME.2003.808824
18. Parer W, Parer J, Holbrook R, Block B. Validity of mathematical methods of quantifying fetal heart rate variability. *Am J Obstetr Gynaecol.* (1985) 153:402–9. doi: 10.1016/0002-9378(85)90078-X
19. Chess GF, La Bell KS, Milne JK, Calaresu FR. Spectral analysis as a diagnostic aid in the management of high-risk pregnancy. *Am J Obstet Gynecol.* (1975) 121:471–4. doi: 10.1016/0002-9378(75)90077-0
20. Van Laar JO, Peters CH, Vullings R, Houterman S, Oei SG. Power spectrum analysis of fetal heart rate variability at near term and post term gestation during active sleep and quiet sleep. *Early Hum Dev.* (2009) 85:795–8. doi: 10.1016/j.earlhumdev.2009.11.001
21. Kwon JY, Park IY, Lim J, Shin JC. Changes in spectral power of fetal heart rate variability in small-for-gestational-age fetuses are associated with fetal sex. *Early Hum Dev.* (2014) 90:9–13. doi: 10.1016/j.earlhumdev.2013.11.005
22. Oppenheimer LW, Lewinsky RM. Power spectral analysis of fetal heart rate. *Baillieres Clin Obstet Gynaecol.* (1994) 8:643–61. doi: 10.1016/S0950-3552(05)80203-2
23. Maeda K, Nagawawa T. Loss of FHR variability diagnosed by frequency analysis. *J Perinat Med.* (2010) 38:197–201. doi: 10.1515/jpm.2010.036
24. Romano M, Iuppariello L, Ponsiglione AM, Improta G, Bifulco P, Cesarelli M. Frequency and time domain analysis of foetal heart rate variability with traditional indexes: a critical survey. *Comput Mathematical Methods Med.* (2016) 2016:1–12. doi: 10.1155/2016/9585431
25. Garabedian C, Jonckheere J, Butruille L, Deruelle P, Storme L, Houfflin-Debarge V. Understanding fetal physiology and second line monitoring during labor. *J Gynecol Obstetr Hum Reprod.* (2017) 46:113–7. doi: 10.1016/j.jogoh.2016.11.005
26. Rassi D, Lewis MJ. Power spectral analysis of the foetal magnetocardiogram. *Physiol Meas.* (1995) 16:111–20. doi: 10.1088/0967-3334/16/2/003
27. van Laar J, Peters CHL, Vullings R, Houterman S, Bergmans JWM, Oei SG. Fetal autonomic response to severe acidaemia during labour. *BJOG Int J Obstetr Gynaecol.* (2010) 117:429–37. doi: 10.1111/j.1471-0528.2009.02456.x
28. Chung DY, Sim YB, Park KT, Yi SH, Shin JC, Kim SP. Spectral analysis of fetal heart rate variability as a predictor of intrapartum fetal distress. *Int J Gynecol Obstetr.* (2001) 73:109–16. doi: 10.1016/S0020-7292(01)00348-4
29. TaskForce. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*. (1996) 93:1043–65.
30. Hayes M. *Statistical Digital Signal Processing and Modelling*. New York, NY: Wiley (1996).
31. Goncalves H, Rocha AP, Ayres-de-Campos D, Bernardes J. Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery. *Med Biol Eng Comput.* (2006) 44:847–55. doi: 10.1007/s11517-006-0105-6
32. Bernardes J, Gonçalves H, Ayres-de-Campos D, Rocha AP. Sex differences in linear and complex fetal heart rate dynamics of normal and acidemic fetuses in the minutes preceding delivery. *J Perinat Med.* (2009) 37:168–76. doi: 10.1515/JPM.2009.024
33. Chudácek V, Spilka J, Lhotská L, Janku P, Koucký M, Huptych M, et al. Assessment of features for automatic CTG analysis based on expert annotation. *Annu Int Conf IEEE Eng Med Biol Soc.* (2011) 2011:6051–4. doi: 10.1109/IEMBS.2011.6091495
34. Gonçalves H, Costa A, Ayres-de-Campos D, Costa-Santos C, Rocha AP, Bernardes J. Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability. *Med Biol Eng Comput.* (2013) 51:665–76. doi: 10.1007/s11517-013-1036-7
35. Zarmehri MN, Castro L, Santos J, Bernardes J, Costa A, Santos CC. On the prediction of foetal acidaemia: a spectral analysis-based approach. *Comput Biol Med.* (2019) 109:235–41. doi: 10.1016/j.compbiomed.2019.04.041
36. Breborowicz G, Moczko J, Gadzinowski J. Analysis of fetal heart rate in frequency domain. In: van Geijn HPCF, editor. *A Critical Appraisal of Fetal Surveillance*. New York, NY: Elsevier (1998). p. 325–32.
37. Sibony O, Fouillot JP, Benaoudia M, Benhalla A, Oury JF, Sureau C, et al. Quantification of the fetal heart rate variability by spectral analysis of fetal well-being and fetal distress. *Eur J Obstet Gynecol Reprod Biol.* (1994) 54:103–8. doi: 10.1016/0028-2243(94)90246-1
38. Kimura Y, Okamura K, Yajima A. Spectral analysis of beat-to-beat intervals of the fetal heart obtained by Doppler ultrasound. *Gynecol Obstet Invest.* (1996) 41:5–9. doi: 10.1159/000292025
39. Chudacek V, Spilka J, Bursa M, Janku P, Hruban L, Huptych M, et al. Open access intrapartum CTG database. *BMC Pregnancy Childbirth.* (2014) 14:16. doi: 10.1186/1471-2393-14-16
40. Ayres-de-Campos D, Arulkumaran S. FIGO consensus guidelines on intrapartum fetal monitoring: physiology of fetal oxygenation and the main goals of intrapartum fetal monitoring. *Int J Gynaecol Obstet.* (2015) 131:5–8. doi: 10.1016/j.ijgo.2015.06.018

41. Executive summary: neonatal encephalopathy and neurologic outcome, second edition. Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Encephalopathy. *Obstet Gynecol.* (2014) 123:896–901. doi: 10.1097/01.AOG.0000445580.65983.d2
42. Nunes I, Ayres-de-Campos D, Ugwumadu A, Amin P, Banfield P, Nicoll A, et al. Central Fetal monitoring with and without computer analysis: a randomized controlled trial. *Obstet Gynecol.* (2017) 129:83–90. doi: 10.1097/AOG.0000000000001799
43. Group IC. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet.* (2017) 389:1719–29. doi: 10.1016/S0140-6736(17)30568-8
44. Siira SM, Ojala TH, Vahlberg TJ, Jalonen JO, Valimaki IA, Rosen KG, et al. Marked fetal acidosis and specific changes in power spectrum analysis of fetal heart rate variability recorded during the last hour of labour. *BJOG.* (2005) 112:418–23. doi: 10.1111/j.1471-0528.2004.00454.x
45. Van Laar JO, Porath MM, Peters CH, Oei SG. Spectral analysis of fetal heart rate variability for fetal surveillance: review of the literature. *Acta Obstet Gynecol Scand.* (2008) 87:300–6. doi: 10.1080/00016340801898950
46. Semmlow J. *Circuits, Signals, and Systems for Bioengineers: A MATLAB-Based Introduction*. London; San Diego, CA; Combridge, MA; Oxford: Elsevier Science (2005).
47. Marôco J. *Análise Estatística com o SPSS Statistics: 7ª edição: ReportNumber, Lda*. Pêro Pinheiro: Report Number (2018).
48. Kwon JY, Park IY, Shin JC, Song J, Tafreshi R, Lim J. Specific change in spectral power of fetal heart rate variability related to fetal acidemia during labor: comparison between preterm and term fetuses. *Early Hum Dev.* (2012) 88:203–7. doi: 10.1016/j.earlhumdev.2011.08.007
49. Rantonen T, Ekholm E, Siira S, Metsälä T, Leino R, Ekblad U, et al. Periodic spectral components of fetal heart rate variability reflect the changes in cord arterial base deficit values: a preliminary report. *Early Hum Dev.* (2001) 60:233–8. doi: 10.1016/S0378-3782(00)00124-9
50. Peters C, Vullings R, Bergmans J, Oei G, Wijn P. The effect of artifact correction on spectral estimates of heart rate variability. *Annu Int Conf IEEE Eng Med Biol Soc.* (2008) 2008:2669–72. doi: 10.1109/IEMBS.2008.4649751
51. Salamalekis E, Hintipas E, Salloum I, Vasios G, Loghis C, Vitoratos N, et al. Computerized analysis of fetal heart rate variability using the matching pursuit technique as an indicator of fetal hypoxia during labor. *J Matern Fetal Neonatal Med.* (2006) 19:165–9. doi: 10.1080/14767050500233290
52. Peters CH, ten Broeke ED, Andriessen P, Vermeulen B, Berendsen RC, Wijn PF, et al. Beat-to-beat detection of fetal heart rate: doppler ultrasound cardiotocography compared to direct ECG cardiotocography in time and frequency domain. *Physiol Meas.* (2004) 25:585–93. doi: 10.1088/0967-3334/25/2/015
53. Madwed J, Albrecht P, Mark R, Cohen R. Low-frequency oscillation in arterial pressure and heart rate: a simple computer model. *Am J Physiol.* (1989) 256:H1573–9. doi: 10.1152/ajpheart.1989.256.6.H1573
54. Pagani M, Montano N, Porta A, Malliani A, Abboud F, Birkett C, et al. Relationship between spectral components of cardiovascular variabilities and direct measures of muscle sympathetic nerve activity in humans. *Circulation.* (1997) 95:1441–8. doi: 10.1161/01.CIR.95.6.1441
55. Randall DC, Brown DR, Raisch RM, Yingling JD, Randall WC. SA nodal parasympathectomy delineates autonomic control of heart rate power spectrum. *Am J Physiol.* (1991) 260 (3 Pt 2):H985–8. doi: 10.1152/ajpheart.1991.260.3.H985
56. Siira S, Ojala T, Ekholm E, Vahlberg T, Blad S, Rosén KG. Change in heart rate variability in relation to a significant ST-event associates with newborn metabolic acidosis. *BJOG.* (2007) 114:819–23. doi: 10.1111/j.1471-0528.2007.01369.x
57. Siira SM, Ojala TH, Vahlberg TJ, Rosén KG, Ekholm EM. Do spectral bands of fetal heart rate variability associate with concomitant fetal scalp pH? *Early Hum Dev.* (2013) 89:739–42. doi: 10.1016/j.earlhumdev.2013.05.007
58. Beer NAMd, Andriessen P, Berendsen RCM, Oei SG, Wijn PFF, Oetomo SB. Customized spectral band analysis compared with conventional Fourier analysis of heart rate variability in neonates. *Physiol Measure.* (2004) 25:1385–95. doi: 10.1088/0967-3334/25/6/004
59. van Laar JO, Peters CH, Houterman S, Wijn PF, Kwee A, Oei SG. Normalized spectral power of fetal heart rate variability is associated with fetal scalp blood pH. *Early Hum Dev.* (2011) 87:259–63. doi: 10.1016/j.earlhumdev.2011.01.028
60. Gonçalves H, Fernandes D, Pinto P, Ayres-de-Campos D, Bernardes J. Simultaneous monitoring of maternal and fetal heart rate variability during labor in relation with fetal gender. *Dev Psychobiol.* (2017) 59:832–9. doi: 10.1002/dev.21554
61. 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) 57:771–9. doi: 10.1109/TBME.2009.2035818
62. Illanes A, Haritopoulos M. Fetal heart rate feature extraction from cardiotocographic recordings through autoregressive model's power spectral- and pole-based analysis. *Annu Int Conf IEEE Eng Med Biol Soc.* (2015) 2015:5842–5. doi: 10.1109/EMBC.2015.7319720
63. Fortrat JO. Inaccurate normal values of heart rate variability spectral analysis in newborn infants. *Am J Cardiol.* (2002) 90:346. doi: 10.1016/S0002-9149(02)02399-8
64. Van Leeuwen P, Geue D, Lange S, Hatzmann W, Grönemeyer D. Changes in the frequency power spectrum of fetal heart rate in the course of pregnancy. *Prenat Diagn.* (2003) 23:909–16. doi: 10.1002/pd.723
65. Otera Y, Morokuma S, Fukushima K, Wake N, Kato K. Correlation between regular mouthing movements and heart rate patterns during non-rapid eye movement periods in normal human fetuses between 32 and 40 weeks of gestation. *Early Hum Dev.* (2013) 89:381–6. doi: 10.1016/j.earlhumdev.2012.12.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 03 August 2021
doi: 10.3389/fped.2021.660476

Distance to Healthy Metabolic and Cardiovascular Dynamics From Fetal Heart Rate Scale-Dependent Features in Pregnant Sheep Model of Human Labor Predicts the Evolution of Acidemia and Cardiovascular Decompensation

Stephane G. Roux[1], *Nicolas B. Garnier*[1]*, *Patrice Abry*[1], *Nathan Gold*[2,3] and *Martin G. Frasch*[4]

[1] Laboratoire de Physique, Université Lyon, Ens de Lyon, Université Claude Bernard, CNRS, Lyon, France, [2] Department of Mathematics and Statistics, York University, Toronto, ON, Canada, [3] Centre for Quantitative Analysis and Modelling, Fields Institute, Toronto, ON, Canada, [4] Department of OBGYN, Center on Human Development and Disability, University of Washington, Seattle, WA, United States

OPEN ACCESS

Edited by:
Beth J. Allison,
Hudson Institute of Medical Research, Australia

Reviewed by:
Ishmael Miguel Inocencio,
Hudson Institute of Medical Research, Australia
Caroline Shaw,
Imperial College London, United Kingdom

***Correspondence:**
Nicolas B. Garnier
nicolas.garnier@ens-lyon.fr

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: 13 February 2021
Accepted: 21 June 2021
Published: 03 August 2021

Citation:
Roux SG, Garnier NB, Abry P, Gold N and Frasch MG (2021) Distance to Healthy Metabolic and Cardiovascular Dynamics From Fetal Heart Rate Scale-Dependent Features in Pregnant Sheep Model of Human Labor Predicts the Evolution of Acidemia and Cardiovascular Decompensation. Front. Pediatr. 9:660476. doi: 10.3389/fped.2021.660476

The overarching goal of the present work is to contribute to the understanding of the relations between fetal heart rate (FHR) temporal dynamics and the well-being of the fetus, notably in terms of predicting the evolution of lactate, pH and cardiovascular decompensation (CVD). It makes uses of an established animal model of human labor, where 14 near-term ovine fetuses subjected to umbilical cord occlusions (UCO) were instrumented to permit regular intermittent measurements of metabolites lactate and base excess, pH, and continuous recording of electrocardiogram (ECG) and systemic arterial blood pressure (to identify CVD) during UCO. ECG-derived FHR was digitized at the sampling rate of 1,000 Hz and resampled to 4 Hz, as used in clinical routine. We focused on four FHR variability features which are tunable to temporal scales of FHR dynamics, robustly computable from FHR sampled at 4 Hz and within short-time sliding windows, hence permitting a time-dependent, or local, analysis of FHR which helps dealing with signal noise. Results show the sensitivity of the proposed features for early detection of CVD, correlation to metabolites and pH, useful for early acidosis detection and the importance of coarse time scales (2.5–8 s) which are not disturbed by the low FHR sampling rate. Further, we introduce the performance of an individualized self-referencing metric of the distance to healthy state, based on a combination of the four features. We demonstrate that this novel metric, applied to clinically available FHR temporal dynamics alone, accurately predicts the time occurrence of CVD which heralds a clinically significant degradation of the fetal health reserve to tolerate the trial of labor.

Keywords: fetal heart rate, animal model, cardiovascular decompensation, distance to healthy state, entropy rate, sliding window analysis, time-scale analysis, sampling rate

1. INTRODUCTION

Monitoring fetal heart rate (FHR) during labor is a common clinical routine worldwide, aiming to asses fetal well-being and ensure safe delivery. The main objective is to decide on timely operative delivery or uterine relaxation to prevent brain injury and adverse outcomes (1). In clinical practice, fetal well-being is assessed by obstetricians principally by visual inspection of cardiotocograms (CTG, bivariate time series of beat-per-minute FHR and uterine activity). The interpretation is guided by a set of rules combining a collection of features, aiming to probe various aspects of the CTG, such as baseline FHR, FHR variability and deceleration shape and timing as well as the relation of the various FHR features to the patterns of uterine activity. One such set of features and rules was defined by the International Federation of Gynecology and Obstetrics (2, 3). Applying such procedure has however been documented as yielding significant inter-and intra-observer variability (4), one of many causes of the failure of the present FHR monitoring to predict fetal brain injury (5–7).

These short-comings in FHR monitoring during labor triggered significant efforts to develop computerized and automated assessment of FHR patterns intrapartum. Beyond the direct computation of the FIGO features themselves [cf., e.g., (8–10)], from digitized CTG usually sampled at 4 Hz in clinical practice, a large variety of features stemming from advanced signal processing and information theory tools has been computed for FHR assessment. These advanced features, however, have not reached performance benchmarks to lead to a consensus in the research and medical communities. These observations leave open a significant number of issues ranging from the choice of relevant FHR features and the construction of decision rules for such features to the assessment of the relationships between FHR time series and fetal well-being. Interested readers are referred to Georgieva et al. (11) (and references therein) for a recent (lack of) consensus overview and the interdisciplinary discussions.

Besides the need for large labeled databases to make machine learning on FHR data effective (12), a recurrent issue is associated with the ground truth being based on pH from the immediate post-birth umbilical cord pH measurements. However, it has been documented that fetal brain injury poorly correlates with measures of acidemia at birth such as pH (7, 11, 13). First, pH is only available after delivery hence when FHR is no longer available. Second, brain compromise due to hypoxia-ischemia can ensue when the fetal cerebral blood flow is persistently reduced, e.g., due to precipitous drop in cerebral perfusion pressure resulting from cardiovascular decompensation (CVD) (7, 14, 15).

In a recent series of experiments, to better assess the relations between FHR, systemic arterial blood pressure (ABP) and fetal health state (including the impact of chronic hypoxia) sheep fetuses were surgically instrumented and subjected to an umbilical cord occlusion (UCO) protocol in Frasch et al. (16), in a well-established animal model of human labor. CVD onset was observed at individually variable times, regardless the presence of chronic hypoxia (7, 16–19). This animal experimental model generated the dataset used in the present study.

Consequently, based on this dataset, the goal of the present work is to assess whether FHR monitoring permits detection of the individual onset of CVD accounting for the presence of chronic hypoxia prior to the onset of UCOs in some fetuses. More particularly, we aimed to assess the sensitivity of FHR temporal dynamics, probed by four scale-dependent features, to CVD, metabolites and pH measurements.

We propose four features which all have the temporal scale of the signal as a parameter. We compute these four quantities from the whole FHR signal to probe its dynamics along the complete experiment. The first quantity measures the average variation of the FHR over the prescribed time scale. The second one measures the FHR variability over the time scale as the standard deviation. The third one is the ratio of the first two and provides a normalized version of the average variation. The fourth one is very similar to Approximate Entropy or Sample Entropy and provides a measure of the information content of the FHR signal at the given time scale. These four quantities can be computed with any signal and give robust results even with the clinically relevant low sampling rate of 4 Hz. This feature choice is also designed to allow computation within short-time time windows, thus permitting to achieve a sliding-window, time-dependent analysis of FHR, which may eventually be exploited to perform real-time FHR monitoring on noisy data.

We show the importance of coarse time scales (2.5–8 s) and construct an individual self-referencing "distance to healthy state" metric based on combination of the four features. We then demonstrate the use of the novel composite distance metric to predict individual CVD from FHR time series alone.

2. MATERIALS: SHEEP ANIMAL MODEL AND UMBILICAL CORD OCCLUSIONS

Fetal sheep model of labor and surgical preparation. The anesthetic and surgical procedures, postoperative care of the animals and the UCO model of labor have been previously described Frasch et al. (16). Briefly, 14 near-term ovine fetuses (123 $\pm$ 2 days gestational age (GA), term = 145 days) of the mixed breed were surgically instrumented. Animal care followed the guidelines of the Canadian Council on Animal Care and was approved by the University of Western Ontario Council on Animal Care.

Polyvinyl catheters were placed in the right and left brachiocephalic arteries, the cephalic vein, and the amniotic cavity. The fetal arterial lines were used for measuring ABP, sampling arterial blood gases, metabolites and cytokines. The fetal venous line was used for administration of fluids and post-operative antibiotics. Stainless steel electrodes were sewn onto the fetal chest to monitor ECG. A polyvinyl catheter was also placed in the maternal femoral vein. Stainless steel electrodes were additionally implanted biparietally on the dura for the recording of electrocorticogram, ECOG, as a measure of summated brain electrical activity [results reported elsewhere (16)]. An inflatable silicon rubber cuff (*in vivo* Metric, Healdsburg, CA) for UCO

induction was placed around the proximal portion of the umbilical cord and secured to the abdominal skin. Once the fetus was returned to the uterus, a catheter was placed in the amniotic fluid cavity. Antibiotics were administered intravenously to the mother (0.2 g of trimethoprim and 1.2 g sulfadoxine, Schering Canada Inc., Pointe-Claire, Canada) and fetus and into the amniotic cavity (1 million IU penicillin G sodium, Pharmaceutical Partners of Canada, Richmond Hill, Canada). Amniotic fluid lost during surgery was replaced with warm saline. The uterus and abdominal wall incisions were sutured in layers and the catheters exteriorized through the maternal flank and secured to the back of the ewe in a plastic pouch. Postoperatively, animals were allowed 4 days to recover prior to experimentation and daily antibiotic administration was continued intravenously to the mother (0.2 g trimethoprim and 1.2 g sulfadoxine), into the fetal vein and the amniotic cavity (1 million IU penicillin G sodium, respectively). Arterial blood was sampled for evaluation of the fetal condition and catheters were flushed with heparinized saline to maintain patency. Animals were 130 ± 1 day GA on the first day of the experimental study.

Umbilical cord occlusion protocol. The experimental protocol has been reported (16, 20, 21). Briefly, all animals were studied over a ∼6 h period. Fetal chronic hypoxia was defined as arterial O2Sat <55% as measured on postoperative days 1–3 and at baseline prior to beginning the UCOs. The first group comprised five fetuses that were also spontaneously hypoxic (n = 5, H/UCO). The second group of fetuses was normoxic (O_2Sat more than 55% before UCOs) (n = 9, N/UCO). As reported, after a 1–2 h baseline control period, the animals underwent mild, moderate, and severe series of repetitive UCOs by graduated inflation of the occluder cuff with a saline solution (16). During the first hour following the baseline period, mild variable FHR decelerations were performed with a partial UCO for 1 min duration every 2.5 min, with the goal of decreasing FHR by ∼30 bpm, corresponding to a ∼50% reduction in umbilical blood flow (22, 23). During the second hour, moderate variable FHR decelerations were performed with increased partial UCO for 1 min duration every 2.5 min with the goal of decreasing FHR by ∼60 bpm, corresponding to a ∼75% reduction in umbilical blood flow. Animals underwent severe variable FHR decelerations with complete UCO, i.e., ∼100% reduction of umbilical blood flow, for 1 min duration every 2.5 min until the targeted fetal arterial pH of <7.00 was detected, at which point the repetitive UCO were terminated. A summary of timings is reported in **Table 1**. These animals were then allowed to recover for 48 h following the last UCO. Fetal arterial blood samples were drawn at baseline, at the end of the first UCO of each series (mild, moderate, severe), and at 20 min intervals (between UCO) throughout each of the UCO series, as well as at 1, 24, and 48 h of recovery. When pH <7.00 was measured, the UCO were stopped and this time point noted as the end of the occlusions. We then obtained the precise pH = 7.00 time point by linear interpolation from this last measured pH value. All blood samples were analyzed for blood gas values, pH, lactate and base excess (BE) with an ABL-725 blood gas analyzer (Radiometer Medical, Copenhagen, Denmark) with temperature corrected to 39.0°C. Plasma from the 4 ml blood samples was frozen and stored for cytokine analysis, reported elsewhere (24). After the 48 h recovery blood sample, the ewe and the fetus were killed by an overdose of barbiturate (30 mg sodium pentobarbital IV, MTC Pharmaceuticals, Cambridge, Canada). A post mortem was carried out during which fetal sex and weight were determined and the location and function of the umbilical occluder were confirmed. The fetal brain was perfusion-fixed and subsequently dissected and processed for later immunohistochemical study (25).

Data acquisition and pre-processing. A computerized data acquisition system was used to record fetal systemic arterial and amniotic pressures and the ECG signal (26). All signals were monitored continuously throughout the experiment. Arterial and amniotic pressures were measured using Statham pressure transducers (P23 ID; Gould Inc., Oxnard, CA). Fetal systemic ABP was determined as the difference between instantaneous values of arterial and amniotic pressures. A PowerLab system was used for data acquisition and analysis (Chart 5 For Windows, ADInstruments Pty Ltd, Castile Hill, Australia). Pressures, ECOG and ECG were recorded and digitized at 1,000 Hz for further study. For ECG, a 60 Hz notch filter was applied. R peaks of ECG were used to derive the heart rate variability (HRV) times series (26). The time series of R-R peak intervals were then uniformly resampled at 4 Hz (26). A representative FHR signal is shown in **Figure 1A**: a visual assessment of the whole FHR trace reveals that FHR variability increases when UCO strength is increased.

Metabolites data (pH, lactate and BE) is obtained by blood sampling performed at specific times during the experiment (vertical dashed lines in **Figure 1B**). In order to have metabolites data at any time, we assume a linear drift between two successive measurements and consequently perform a linear interpolation between two measurements times. We thus obtain a piece-wise linear time series sampled at 4 Hz, depicted in **Figure 1B** as a black curve. Using this interpolated data, as noted above, the time t_{pH} when pH = 7.00 is computed in each fetus as indicated in **Table 1**.

Fetal cardiovascular decompensation (CVD). CVD has been reported in detail in Frasch et al. (15, 16) and Gold et al. (27). The visual representation of CVD can be found in these publications, e.g., in the Figure 2 in Frasch et al. (16). The reader can readily observe the pronounced pathological hypotensive responses to the UCO-triggered FHR decelerations during the CVD. This behavior is in stark contrast to the normally observed ABP increases during the occlusions which compensate the hypotension caused by the FHR decelerations. As we reported, once this pattern conversion from hypertensive to hypotensive responses occurs, it persists until the UCOs are stopped. Its effects are also seen directly in the brain electrical activity (15, 16). It is hence easy to reliably visually identify the timing of the onset of CVD in each recording. Consequently, during UCOs, by expert visual inspection, we noted the individual time point t_{CVD} at which three successive hypotensive ABP responses to UCO-triggered FHR decelerations occurred. Quantitatively, with hypotensive ABP response we refer to the failure of ABP to rise during UCO-triggered FHR deceleration above the preceding baseline value when compared to the average ABP rise during the

TABLE 1 | Individual onset times for each UCO regime (mild, moderate, severe and recovery, colored green, magenta, red and white in **Figure 1**), counted from the first UCO.

		UCO start time				pH = 7.00 time
	Animal (ID)	Mild (hh:mm)	Moderate (hh:mm)	Severe (hh:mm)	Recovery (hh:mm)	t_{pH} (hh:mm)
	8003	(01:14)	00:57 (02:11)	02:00 (03:14)	02:07 (03:21)	02:07 (03:21)
	473351	(NaN)	00:00 (04:08)	01:08 (05:16)	02:33 (06:41)	01:44 (05:52)
Hypoxic	473376	(02:53)	00:55 (03:48)	01:51 (04:44)	02:54 (05:47)	01:58 (05:51)
	473726	(02:09)	01:00 (03:09)	01:55 (04:04)	03:17 (05:26)	03:16 (05:25)
	473362	(02:08)	01:01 (03:09)	01:54 (04:02)	02:20 (04:38)	02:31 (04:39)
	473352	(NaN)	00:00 (03:59)	01:00 (04:59)	01:46 (05:45)	01:39 (05:38)
	5054	(01:31)	00:56 (02:27)	02:00 (03:31)	03:51 (05:22)	03:51 (05:22)
	461060	(02:59)	00:54 (03:53)	01:59 (04:58)	03:30 (06:29)	03:30 (06:29)
	5060	(01:09)	00:57 (02:06)	01:59 (03:08)	02:58 (04:07)	02:53 (04:02)
Normoxic	473360	(02:11)	01:05 (03:16)	02:02 (04:13)	03:59 (06:10)	03:59 (06:10)
	473378	(03:17)	00:58 (04:15)	01:53 (05:10)	02:31 (05:48)	02:28 (05:45)
	473727	(01:38)	01:05 (02:43)	02:02 (03:40)	04:10 (05:48)	03:40 (05:18)
	473377	(02:28)	01:04 (03:32)	02:03 (04:31)	04:04 (06:32)	03:59 (06:27)
	473361	(01:56)	01:03 (02:59)	02:05 (04:01)	03:26 (05:22)	03:29 (05:25)

*Values in parenthesis are times counted from the beginning of the recording as represented on the time-axis of **Figures 1**, **8**, **9**, **10**. For animals 473351 and 473352, the first UCO had a moderate effect (of decreasing FHR by about 60 bpm), so phase names have been shifted accordingly.*

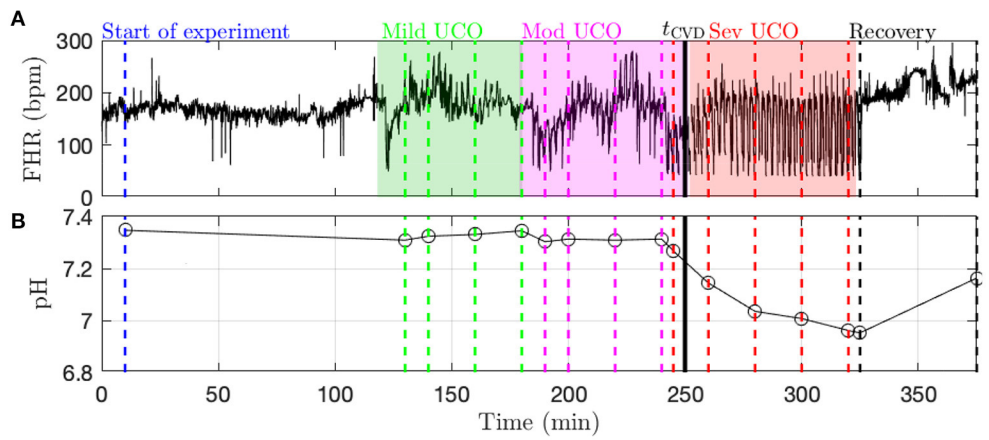

FIGURE 1 | Typical data recorded in the experiment (here, animal 473726). **(A)** FHR resampled at 4 Hz. The color indicates the strength of UCO during the experiment: blue and black for no UCO (baseline and recovery), green for mild UCO, magenta for moderate UCO and red for severe UCO. **(B)** Fetal arterial pH values during the experiment. The pH (as well as the other blood measurements) is obtained at specific time points, indicated by the vertical lines; the open black circles correspond to the actual measurements from blood sampling and the black lines correspond to a linear interpolation.

UCO series prior to the CVD. We refer to this animal-specific time point as the ABP sentinel corresponding to the timing of CVD. As an illustration, t_{CVD} is reported in **Figure 1B** as a vertical black line.

3. METHODS: TIME SCALE-DEPENDENT FEATURES

3.1. Sliding Window Analysis

Analysis of FHR and metabolites data are performed in sliding time-windows of size $T = 20$ min. The time-windows are shifted by $dT = 5$ min, thus implying a $T - dT = 15$ min (75%) overlap. The k-th time-window thus corresponds to time ranging in $[kdT, kdT + T]$. This sliding window analysis permits the assessment of the temporal evolution of cardiovascular responses to changes in UCO strength.

3.2. Scale Dependent Features

Using FHR, x_t, four quantities, whose definitions rely on the choice of a time scale τ, are computed, for each time-window k: increment mean $m_k(\tau)$, increment standard deviation $\sigma_k(\tau)$, the corresponding Student ratio $R_k(\tau)$, and the entropy rate

$h_k(\tau)$. The first one, $m_k(\tau)$, measures the average change of the FHR signal over the time lag τ. It can be pictured as the derivative of the signal on the time-scale τ, averaged in the time-window k. The second one, $\sigma_k(\tau)$, is the standard deviation of the signal estimated on chunks of the signal of duration τ, and then averaged in the time-window k. The third one, $R_k(\tau)$, is a normalized version of the first quantity $m_k(\tau)$: the average variation is now expressed in standard deviation units, before being averaged along the time-window k. The last one, $h_k(\tau)$, in a measure of the information or complexity of the signal at scale τ, similar to Approximate Entropy or Sample Entropy, estimated at scale τ over the time-window k. As described below, the first three quantities are averages over the time-window k of dynamical quantities $m_t(\tau), \sigma_t(\tau), R_t(\tau)$ defined at scale τ; these quantities, along with FHR, are shown for illustration purposes in **Figure 2** for arbitrarily chosen 20-min window, time scale τ and animal.

3.2.1. Mean Variation (or Local Trend) at Time-Scale τ

For all t in window k, the average increment over a time scale τ is computed as:

$$m_t(\tau) = \frac{1}{\tau} \sum_{i=t-\tau+1}^{t} (x_i - x_{t-\tau}) = \frac{1}{\tau} \sum_{i=t-\tau+1}^{t} x_i - x_{t-\tau} \,. \quad (1)$$

These $m_t(\tau)$ are then averaged across window k, for all non-overlapping time-intervals $[(j-1)\tau; j\tau]$:

$$m_k(\tau) = \frac{1}{\lfloor T/\tau \rfloor} \sum_{j=1}^{\lfloor T/\tau \rfloor} m_{kT+j\tau}(\tau) \,, \quad (2)$$

where $\lfloor T/\tau \rfloor$, the floor of the fraction T/τ, indicates the number of time-intervals of size τ available in the time-window of size T. An illustration of the methodology is given in **Figure 2B** for the window $k = 0$: values of $m_t(\tau)$ are depicted in black, and the single value $m_k(\tau)$ is represented in red.

The quantity $m_k(\tau)$ measures the average variation — either an increase or a decrease — of FHR on the time scale τ. The average is indeed a double average: first over all time scales smaller than τ, according to Equation 1, and second over all available intervals available in the kth time-window of size T=20 min, according to Equation (2). $m_k(\tau)$ can also been interpreted as the averaged derivative of the signal after a low-pass filtering using a finite impulse response with cut-off frequency $1/\tau$.

3.2.2. Standard Deviation at Time-Scale τ

Given a time-interval $[t-\tau; t]$, we define the variance of the set of increments $\{(x_{t-i} - x_{t-\tau}),\ t-\tau < i \leq t\}$. This indeed is nothing but the variance of x_t, computed over the set of values in the time-interval $[t-\tau; t]$:

$$\sigma_t^2(\tau) = \frac{1}{\tau} \sum_{i=t-\tau+1}^{t} x_i^2 - \left(\frac{1}{\tau} \sum_{i=t-\tau+1}^{t} x_i \right)^2 . \quad (3)$$

We then average its square root over the $\lfloor T/\tau \rfloor$ non-overlapping time intervals of size τ available in the kth time-window of size T:

$$\sigma_k(\tau) = \frac{1}{\lfloor T/\tau \rfloor} \sum_{j=1}^{\lfloor T/\tau \rfloor} \sigma_{kT+j\tau}(\tau) \,. \quad (4)$$

This quantity measures the average — in the kth time-window of size T — amplitude of the fluctuations of x_t over τ consecutive points. The methodology is illustrated in **Figure 2C**.

3.2.3. Normalized Local Trend at Time-Scale τ

The Student ratio, or normalized local trend, at time-scale τ, is defined for each time interval $[t-\tau; t]$, as:

$$R_t(\tau) = \frac{m_t(\tau)}{\sigma_t(\tau)} \,. \quad (5)$$

It is averaged across all available non-overlapping intervals in the kth time-window:

$$R_k(\tau) = \frac{1}{\lfloor T/\tau \rfloor} \sum_{j=1}^{\lfloor T/\tau \rfloor} R_{kT+j\tau}(\tau) \,. \quad (6)$$

This quantity, up to a factor $\sqrt{\tau}$, would correspond to a random variable drawn from the distribution of the t-value if the data x_t were independently drawn from a Gaussian distribution. It can be interpreted as the average variation over a time step τ, normalized by the local standard deviation; as such, it provides a normalized measure of the trend of the signal x_t to depart from its expected value when observed across a duration τ.

3.2.4. Entropy Rate at Time-Scale τ

One commonly used feature in heart rate analysis, both for adults and fetuses, is sample entropy (SampEn) (28–30), an elaboration on approximate entropy (ApEn) (31, 32). It was shown recently that the entropy rate provides a related tool to probe FHR with better performance than ApEn or SampEn to detect acidosis (33–35).

The entropy rate of order 1 in the kth time-window at time-scale τ is defined as:

$$h_k(\tau) = H(x_t, x_{t-\tau}) - H(x_t) \,, \quad (7)$$

where

$$H(\vec{x}) = -\int p(\vec{x}) \ln p(\vec{x}) d\vec{x} \,, \quad (8)$$

denotes the Shannon entropy (36) of either a vector $\vec{x} = (x_t, x_{t-\tau})$ or a scalar $\vec{x} = x_t$. $h_k(\tau)$ is computed using all the pairs of points $(x_t, x_{t-\tau})$ available in the k-th time-window, and following Theiler's prescription (34) to avoid spurious correlation.

$h_k(\tau)$ measures the extra information conveyed by the vector $(x_t, x_{t-\tau})$ when (x_t) is known, or in other words, the extra information given by the knowledge of the signal at an earlier time $t-\tau$. The entropy rate probes the dynamics of the signal, and to better focus on this dynamical aspect, we compute it on the normalized signal $(x_t - \langle x_t \rangle)/\sqrt{\langle (x_t - \langle x_t \rangle)^2 \rangle}$, where $\langle . \rangle$ stands for the time average on the window of size T.

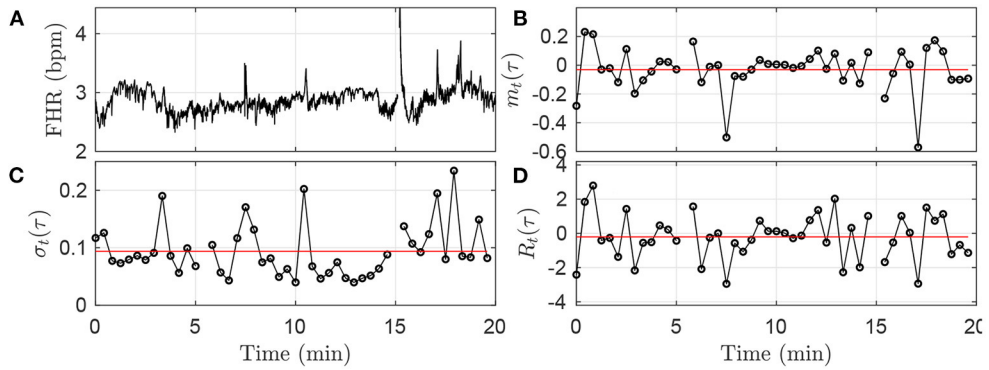

FIGURE 2 | Illustration of the methodology in the first time-window [0; T] of size $T = 20$ min, using the fixed time scale $\tau = 25$s. **(A)**: FHR. **(B)**: $m_t(\tau)$. **(C)**: $\sigma_t(\tau)$. **(D)**: $R_t(\tau)$. The black circles in **(B–D)** correspond each to a value obtained in a time-interval of size $\tau = 25$ s, according to Equations (1), (3), (5). The horizontal red lines in **(B–D)** indicate the values $m_{k=1}(\tau)$, $\sigma_{k=1}(\tau)$ and $R_{k=1}(\tau)$ obtained after averaging all the black circles, i.e., over all available time-intervals of size τ in the time-window, according to Equations (2), (4), (6).

4. RESULTS AND DISCUSSION: FEATURES, TIME-SCALES AND DISTANCE TO HEALTHY STATE

In section 4.1, the four features — computed in overlapping time-windows — evolution in time are firstly presented and studied with respect to their relations to UCO strength. Because these features are computed at a given time-scale τ, they offer a description of the FHR dynamics at this time-scale. We thus explore the correlation between the features at a given time-scale τ and the measured values of the metabolites — including the pH. This global analysis, presented in section 4.2, is performed using all available time-windows and all available animals. We then reduce the dimensionality of the analysis by averaging results over the long-term time-scales, as defined and presented in section 4.3. This allows us to examine more clearly how the features evolve jointly with the UCO strength for the entire cohort, while quantifying the variability between animals. We then examine quantitatively in section 4.4 how these long-term features correlate with metabolites. We then combine them in an appropriately normalized vector; we are then able to describe the large variability across the subjects in the population as the variability of this vector in the early stages of the experiments. This allows us to define a measure of the degradation of the health state of an animal as the distance from healthy state. Finally, we propose in section 4.5 to use this "individual" distance as a novel indicator — or sentinel — to alert for the degradation of the health status due to CVD. We also show that this indicator/sentinel matches very well with pH measurements.

4.1. Features and UCO Strength

We first examine on a single animal how the four FHR features evolve throughout an experiment, depending on the time-scale τ. The values obtained in the k-th time-window $[kdT; kdT + T]$ are assigned to the date $t_k = kdT + T/2$ at the center of the time-window. The dynamical evolutions of $m_k(\tau)$, $\sigma_k(\tau)$, $R_k(\tau)$ and $h_k(t, \tau)$ are depicted in **Figure 3** for a large band of time scales τ.

Such a time-scale representation reveals qualitatively that when the UCO strength is increased, $m_k(\tau)$, $R_k(\tau)$ and $h_k(\tau)$ decrease along time, while $\sigma_k(\tau)$ increases along time. This agrees with the previous studies where the decrease of the entropy rate $h_k(\tau)$ was associated with fetal acidosis (33–35).

Qualitatively, although the four features barely evolve in time for smaller values of τ (below 2 s, bottom of the images in **Figure 3**), a noticeable time evolution can be observed for large values of τ and especially in the severe UCO regime. To better observe the dependence of the four features on the scale τ, we plot in **Figure 4** their evolution with τ for the time points when blood sampling was performed. **Figure 4** therefore presents the evolution of the four features along the vertical color lines indicated in the images of **Figure 3**.

We observe in **Figure 4** that the evolution of $m_k(\tau)$ is rather linear in τ, but the slope depends on the time, and hence on the UCO level. We observe almost no evolution of $R(\tau)$ with τ, but the value of $R(\tau)$ depends on time, so on the UCO level. On the contrary, both $\sigma(\tau)$ and the entropy rate $h(\tau)$ present a distinct change of their evolution with τ below and above $\tau = 2.5$ s, which emphasizes the distinction between short (< 2.5 s) and large (> 2.5 s) time scales, in accordance with previous literature (26, 37, 38). We use this information on the time scales as follows.

4.2. FHR Features, Arterial Metabolites and pH

We now examine, for a fixed time-scale τ, how the features relate to the health state of the animal, as described by the metabolites and pH. To do so, we use all time-windows of size T on one side, and interpolated metabolites data on the other side. We compute the correlation between any of the four features (for a fixed τ) and any of the biochemical measurements, by averaging over all time-windows (average over k) and over

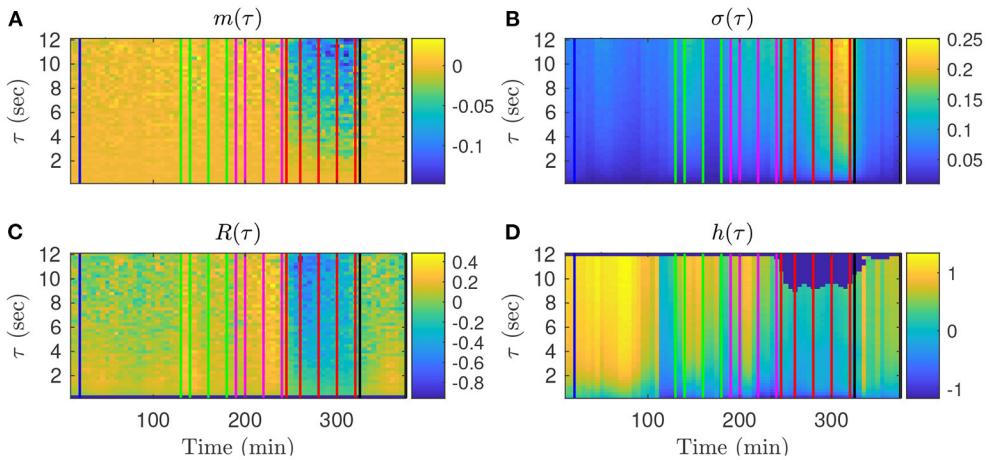

FIGURE 3 | Representation of the time evolutions of $m_k(\tau)$ **(A)**, $\sigma_k(\tau)$ **(B)**, $R_k(\tau)$ **(C)**, $h_k(\tau)$ **(D)**, depending on the scale τ for animal 473726. The time in abscissa is $kdT + T/2$, the location of the kth time-window of size T where the quantity is computed, and the ordinate represents the scale τ. Vertical color lines indicates the times at which blood sampling was performed (same color code as in **Figure 1**: green in the mild UCO regime, magenta in the moderate UCO regime, and red in the severe UCO regime). In the severe UCO regime and for larger time scales τ, stronger variations are observed.

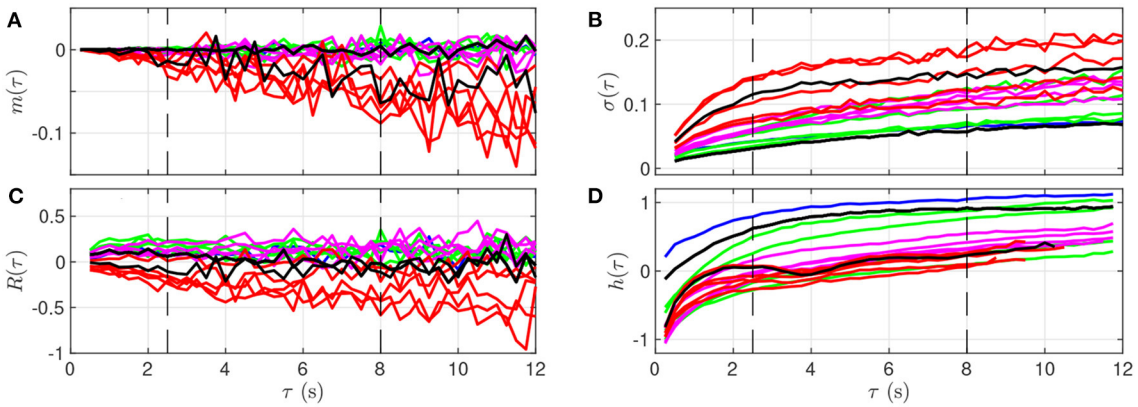

FIGURE 4 | Quantitative representation of the evolution of $m(\tau)$ **(A)**, $\sigma(\tau)$ **(B)**, $R(\tau)$ **(C)**, and $h(\tau)$ **(D)** over the time scale τ for a single animal. The data represented here is extracted from **Figure 3**: each curve corresponds to a time-window of size T for which a fetal arterial blood sample was taken. The color of the curve represents the corresponding UCO level, with the same color code as in **Figures 1**, **3**: blue is the baseline prior to any UCO, green in the mild UCO regime, magenta in the moderate UCO regime, red in the severe UCO regime, and then black in the recovery regime (after UCO). Vertical black dashed lines indicate the time-scales 2.5 and 8 s.

all animals. Results are plotted in **Figure 5** as a function of the scale τ.

As suggested by **Figure 4** and confirmed by **Figure 5**, we can isolate two bands of time scales: shorter scales $\tau < 2.5$s (i.e., high frequency band, above 0.4 Hz, short term time scales, labeled ST) and larger ones $\tau > 2.5$s (low frequency band, below 0.4 Hz, long term time scales, labeled LT).

For any of the four features and any of the three biochemical measurements, the correlation in the range $[2.5 - 8]$ s is not only the largest—in absolute value—but also the most stable: it fluctuates less and does not depend much onτ. Above 8 s, all correlations decrease in absolute value, which may be attributed in part to poorer statistics: the number $\lfloor T/\tau \rfloor$ of available time-intervals of size τ in a time-window of size T decreases, which impacts the averages, see, e.g., Equation 2). As a consequence, we choose in the following to restrict the long term (LT) range to $\tau \leq 8$ s in order to have enough statistical power.

4.3. Long-Term Scales Averaged FHR Features

For the sake of simplicity, we now eliminate the dependencies of our features on τ and focus on the LT range. To do so, we compute the area under the curve (AUC) of our four FHR features in the range $2.5 < \tau < 8$ s. For a given time-window indexed by k, we compute:

$$m_k^{\mathrm{LT}} = \sum_{\tau=2.5\mathrm{s}}^{\tau=8\mathrm{s}} m_k(\tau) \tag{9}$$

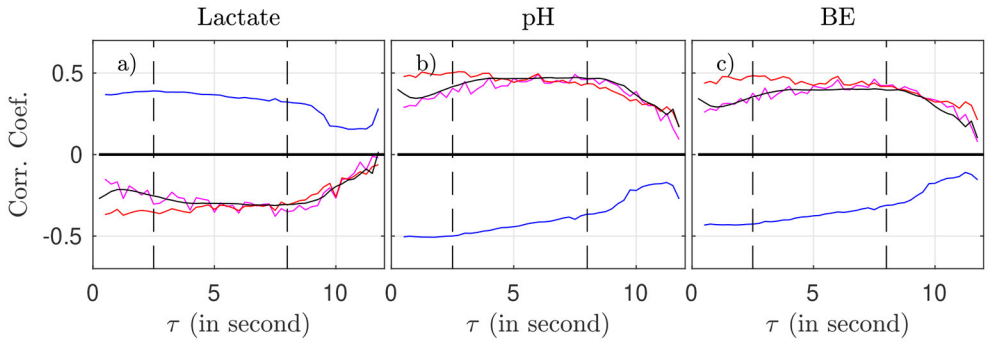

FIGURE 5 | Correlation coefficient between the biochemical measurements (**a**: lactate, **b**: pH, **c**: BE) and the four FHR features: $m_k(\tau)$ (magenta), $\sigma_k(\tau)$ (blue), $R_k(\tau)$ (red), and entropy rate $h_k(\tau)$ (black), as function of the scale τ.

and we define accordingly $\sigma_k^{\rm LT}$, $R_k^{\rm LT}$ and $h_k^{\rm LT}$. These features depend only on time, via the index k of the time-window in which they are computed.

Time evolutions of these four LT features are depicted in **Figure 6** for the complete set of 14 animals. For some animals, there may be missing data due to experimental conditions, and hence there may be less consecutive time-windows of size T available than expected in a given UCO region; in that situation, we have then chosen to assign the dark blue color (arbitrary) for the quantity—see, e.g., the second line (a hypoxic animal), where no data is available in the mild UCO region, and only 4 windows are available in the severe UCO region.

Using the first column (on the left of the vertical green line) of each subfigure as a reference, we observe that every quantity evolves as the UCO strength is increased. Although very few changes are observed in the mild UCO region, much larger variations are observed in the severe UCO region. After the stopping of UCOs (on the right of the vertical black line), we observe that the four features seem to regain their original value, which we interpret as indicating the recovery of the animal, typically after 1 window of size T, so typically 20 min after the end of UCOs.

4.4. Distance to Healthy State, Metabolites and pH

We now explore how our four FHR features relate to the metabolites' levels, and especially to the pH value, which is a widely used indicator of fetal well-being. We report in **Table 2** the correlation coefficient between each of the four features $m_{\rm LT}$, $\sigma_{\rm LT}$, $R_{\rm LT}$ and $h_{\rm LT}$ on one hand, and the three biochemical measurements pH, BE and lactate on the other hand. To increase the statistical power, we use all available time-windows of size T and so all linearly interpolated values of the three biochemical measurements.

We observe that the four FHR features correlate well with the pH and BE, while the correlation with the lactate is smaller. All features but $\sigma^{\rm LT}$ — the LT amplitude of fluctuations — have a correlation coefficient with pH that is at least 0.50, and a correlation coefficient with BE that is at least 0.43. This is interesting, as $R_{\rm LT}$ appears strongly correlated with $m_{\rm LT}$ while relatively uncorrelated with $\sigma_{\rm LT}$.

We believe that each of the four FHR features contributes a particular piece of information about FHR and we therefore aggregate them as follows. For a single animal and a single time-window indexed by k, we consider the vector

$$\vec{u}_k = \left(\frac{m_k^{\rm LT}}{m_{\rm RMS}^{\rm LT}}, \frac{\sigma_k^{\rm LT}}{\sigma_{\rm RMS}^{\rm LT}}, \frac{R_k^{\rm LT}}{R_{\rm RMS}^{\rm LT}}, \frac{h_k^{\rm LT}}{h_{\rm RMS}^{\rm LT}} \right) , \tag{10}$$

where each component is normalized by its standard deviation computed over all animals and over all available time-windows of size T. The four values $(m_{\rm LT}^{\rm RMS}, \sigma_{\rm LT}^{\rm RMS}, R_{\rm LT}^{\rm RMS}, h_{\rm LT}^{\rm RMS})$ used for this normalization are hence the same for all animals and all time-windows; they are reproduced in the third line of **Table 3**.

For a given animal and for a given time-window indexed by k, we use the $\mathcal{L}^2$ norm in $\mathbb{R}^4$ to project any vector $\vec{u}_k$ into a positive real number $\|\vec{u}_k\|$ as follows. For each animal, we assume it is in a healthy condition when the experiment is started (so the FHR is fluctuating around the baseline) and we use the first time window of size T as a reference. We thus define the distance between $\vec{u}_k$ which describes the state in the k-th time-window and $\vec{u}_0$ which describes the state in the first time-window $[0; T]$:

$$D_k = \|\vec{u}_k - \vec{u}_0\| \ . \tag{11}$$

We interpret this distance D_k for a single animal as a measure of the deviation from the animal's "healthy" state during the experiment.

We report in **Table 3** global statistics — obtained by considering all animals — of the four FHR features used as the four components of the vector $\vec{u}_t$.

The third line of **Table 3** reports the values $m_{\rm RMS}^{\rm LT}, \sigma_{\rm RMS}^{\rm LT}, R_{\rm RMS}^{\rm LT}$ and $h_{\rm RMS}^{\rm LT}$ used to normalize the vector $\vec{u}_k$. Their amplitude is notably different, and the normalization is necessary to ensure that each component of $\vec{u}_k$ contributes equally to its norm $||\vec{u}_k||$. Whereas, this normalization uses all available data (using all times and all animals at once), it is important to stress that we have accounted for the large variability from one animal to another by defining D_k with a reference relative to the very animal

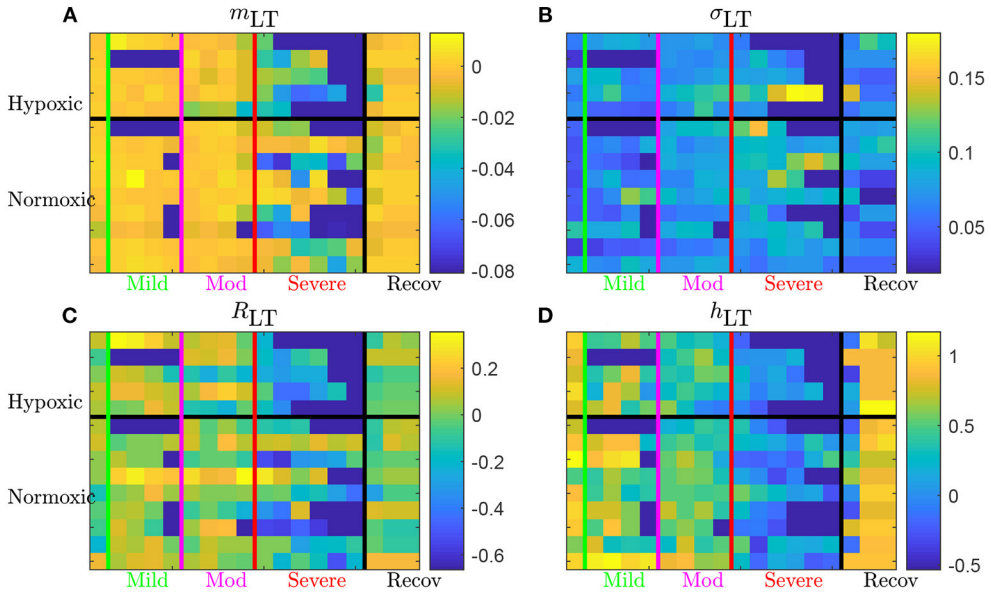

FIGURE 6 | Long term AUC of the four FHR features m_{LT} **(A)**, σ_{LT} **(B)**, R_{LT} **(C)**, and h_{LT} **(D)** for all 14 animals. For a given quantity, each line represents an animal (ordered from top to bottom as in **Table 1**); chronically hypoxic ones are above and normoxic ones are below. Each column represent a time-window of size T and time is increasing from left to right. The region of mild UCO starts at the vertical green line and lasts for four windows, up to the vertical magenta line, followed by 4 windows in the moderate UCO region, and then up to six windows in the severe UCO region, and up to three windows in the recovery region.

TABLE 2 | Correlation coefficients between the four individual features, their vectorial combinations, and the three measurements pH, BE, and Lactate.

	m_{LT}	σ_{LT}	R_{LT}	h_{LT}	$\|\|\vec{u}\|\|$	D	pH	BE	Lactate
m_{LT}	1.00	–0.51	0.77	0.60	0.29	–0.87	0.53	0.48	–0.36
σ_{LT}	–0.51	1.00	–0.19	–0.43	–0.35	0.61	–0.42	–0.36	0.35
R_{LT}	0.77	–0.19	1.00	0.42	0.14	–0.63	0.50	0.48	–0.35
h_{LT}	0.60	–0.43	0.42	1.00	0.89	-0.76	0.50	0.43	–0.32
norm $\|\|\vec{u}\|\|$	0.29	–0.35	0.14	0.89	1.00	–0.50	0.35	0.29	–0.21
distance D	–0.87	0.61	–0.63	–0.76	–0.50	1.00	–0.61	–0.53	0.44
pH	0.53	–0.42	0.50	0.50	0.35	–0.61	1.00	0.95	–0.77
BE	0.48	–0.36	0.48	0.43	0.29	–0.53	0.95	1.00	–0.72
Lactate	–0.36	0.35	–0.35	–0.33	–0.21	0.44	–0.77	–0.72	1.00

Data from all 14 animals and all available time-windows were used. Colored values indicate the quantity which correlates better with the biochemical measurement; it is always the distance D.

TABLE 3 | Means and standard deviations (std) of the four FHR features over the population of 14 animals.

	m_{LT}	σ_{LT}	R_{LT}	h_{LT}
Mean, over animals and over k	–0.0067	0.0688	–0.0262	0.5957
Mean, over animals, fixed $k = 0$	–0.0009	0.0579	0.0128	0.8811
Std, over animals and over k	0.0155	0.0282	0.1630	0.4127
Std, over animals, fixed $k = 0$	0.0040	0.0221	0.1082	0.1970

First and third lines: averages over animals and over time-windows (k). Second and fourth lines: averages over animals, using the first ($k = 0$) time-window $[0; T]$ only.

under consideration. The variability of the reference point can be seen in the fourth line of **Table 3**: it accounts for a large part of the RMS values used in the normalization. Comparing the first two lines of **Table 3** brings an additional observation leading to the same conclusion: the position of the healthy state $\vec{u}_0$ is on average over the animals (second line of the table) sensibly different from the position of $\vec{u}_k$ averaged over all animals and all times (first line of table). Using D_k instead of $||\vec{u}_k||$ removes a large part of the inter-animal variability and definitely improves the relevance of the distance, as measured by the correlation with the metabolites, see **Table 2**.

We present in **Figure 7** the 14 trajectories of the vector $\vec{u}(k)$ in its phase space, for the complete cohort. $\vec{u}_k$ has 4 coordinates so there are 6 different projections in a plane defined by two variables. Each subplot in **Figure 7** corresponds to one of these possible projections. Along each trajectory, the color changes to indicate the interpolated pH value. Although the trajectories wander in a large region of the phase space, their color-coding seem to only depend on the distance from the origin: blue (larger pH) close to the origin, and orange or red (lower pH) outside of the circle defined by $D = 2$. During an experiment, the UCO's strength increases and, as a consequence, the pH decreases. We observe that the distance D_k appears to increase concomitantly, and more precisely we observe its correlation with the pH value. The correlation coefficients between the distance D and the biochemical measurements, computed over all animals, are reported in **Table 2** (gray-colored cells). We observe that among all FHR features we have computed, the

distance D is the one that is the most correlated with pH, as well as with the other metabolites. This confirms that using all four FHR features simultaneously—by considering the vector $\vec{u}$—not only mitigates the various evolutions of single features with the metabolite value but also aggregates their correlations.

4.5. Distance to Healthy State as a New Sentinel for CVD

To better illustrate the relation between the dynamical features—especially the distance D_k—on one hand, and the health status as assessed by the metabolites—especially the pH—and blood pressure responses to UCOs on the other hand, we examine in detail in **Figures 8**, **9** how these are co-evolving for each individual animal. From now on, we discard any indication of the UCO level.

Figure 8 presents jointly the time-trace of the FHR signal (**Figures 8a,e**) and the evolution of the distance D (**Figures 8b,f**), which we color-code using the pH value, as in **Figure 7** for two typical normoxic animals. We also present two projections of the trajectory of $\vec{u}_k$ in its phase space in order to illustrate the evolution of each of the four quantities.

Because the vector $\vec{u}_k$ has been carefully normalized, and appropriately centered to define the distance D_k, this last quantity has no dimension and can be compared to absolute values. The particular value $D = 1$ (gray circle) defines the standard deviation range in a healthy situation and the value $D = 2$ (black circle) corresponds to variations with an amplitude of 2 standard deviations. Looking at the trajectory projections in **Figures 8**, **9**, we see that during the early stages of the experiments the trajectory remains close to the origin, hence the distance D remains small, albeit fluctuating, and the pH value remains close to its normal value (bluish color, indicating a pH close to 7.4). More interestingly, we see that when the pH decreases down to 7.2 (greenish color), the trajectory usually reaches the black circle, hence the distance increases up to 2. Finally, we observe that when the trajectory is outside the black circle, hence $D > 2$, the pH has low values but more importantly, values of pH$\leq$7.00 (orange to red color) are only observed on the trajectory much later after the trajectory wandered outside the black circle.

The very same observations can be made for hypoxic animals, see **Figure 9** for two examples.

We now examine quantitatively for each animal the time evolution of the distance from its own reference healthy state. In every time-window $[kdT, kdT + T]$ of size T indexed by k, we have a value D_k which we assign to time $t = kdT + T/2$; we plot the time-evolution of the distance in **Figures 8b,f**, **9b,f** with a color that indicates the pH value. This allows for the following interesting observations. First, before the occurrence of UCO, the distance fluctuates with a typical standard deviation of 1. This confirms that the normalization step is valid, albeit it uses values averaged over all animals and all available time-windows. Second, we see that the distance D is substantially larger when UCOs are performed, and more precisely, we see that D increases as the UCO strength is increased or UCOs are being applied to the animal with the same strength but for a longer period of time. As such, the distance D seems to be a good representation of the health condition of the animal. More interestingly, we showed (**Table 2**) that D is highly correlated with the pH value throughout the complete experiment, but we now observe that a large value of D, above 2, proves to be a very good indicator of a low pH value signaling an acidemia with pH < 7.2.

To further test the ability of the distance D to alert on the fetal condition, we now try to relate the large values of the distance D to the onset of the fetal CVD, i.e., failure of the fetus to mount a hypertensive arterial blood pressure response to UCOs and the UCO-induced FHR decelerations, a prerequisite to maintaining an adequate cerebral perfusion pressure. To do so, we use t_{CVD} as the reference time when CVD occurs (vertical red line in **Figures 8b,f**, **9b,f**), which offers a valuable benchmark for an early detection of hypotensive blood pressure response.

t_{CVD} appears on group average for a pH of 7.20 and 60 min prior to $\mathrm{pH_{nadir}}$ of <7.00, but shows a considerable inter-individual spread. To quantify the hypotensive behavior at t_{CVD}, we report the individual pressure differential ΔABP at t_{CVD} in **Table 4** following the same approach as reported in Table 1 of Frasch et al. (16). Here, $\Delta\mathrm{ABP}{=}\mathrm{ABP}_{max} - \langle\mathrm{ABP}\rangle$ is the difference between ABP_{max}, the maximal ABP during a UCO, and $\langle\mathrm{ABP}\rangle$, the mean ABP between UCOs. The average of ΔABP is 4±6 mmHg for hypoxic fetuses and 4±8 mmHg for normoxic fetuses. Overall, we see no difference in ΔABP between hypoxic and normoxic groups ($p = 0.97$). The corresponding drop in ΔABP during the CVD period compared to the preceding UCO period is -19 [-24; -1] mmHg, i.e., during CVD, fetuses failed to mount hypertensive response to UCO-triggered FHR decelerations with a median drop of 19 mmHg compared to ΔABP preceding the CVD. These values clearly indicate the pathological hypotensive responses of the sheep fetuses during the UCOs at t_{CVD} and onward until the end of the UCOs. The noted inter-individual variability in ΔABP values is subject of ongoing research.

Here we see on phase-space projections in **Figures 8c,d,g,h**, **9c,d,g,h** that the criterion $D \geq 2$ offers a similar early alert on the deterioration of the animal condition with regard to CVD timing. Looking at either the phase space representation or the time traces of the distance D, we see that this quantity evolves continuously in time, on typical time-scales larger than 20 min, the duration we have chosen to compute our quantity. The distance increases over the duration of the experiment and one can easily measure the time t_D at which D crosses the value $D = 2$ (red circle, or horizontal red line in **Figures 8**, **9**). Unfortunately, the distance D is very sensitive and it can be seen on the examples that it is possible for D to reach values larger than 2 early in the experiment. To overcome these events—and, hence, to make our new sentinel less sensitive—we arbitrarily adjust our criteria and require $D_k > 2.5$ for at least 3 consecutive time-windows, so for a long enough duration of about 40 min. **Table 1** presents the various timings corresponding to the various UCO regimes for each animal, together with an estimate of the pH nadir time, while **Table 4** presents a summary of our findings, together with the CVD time (ABP sentinel), the two of them appearing before $pH \leq 7.00$.

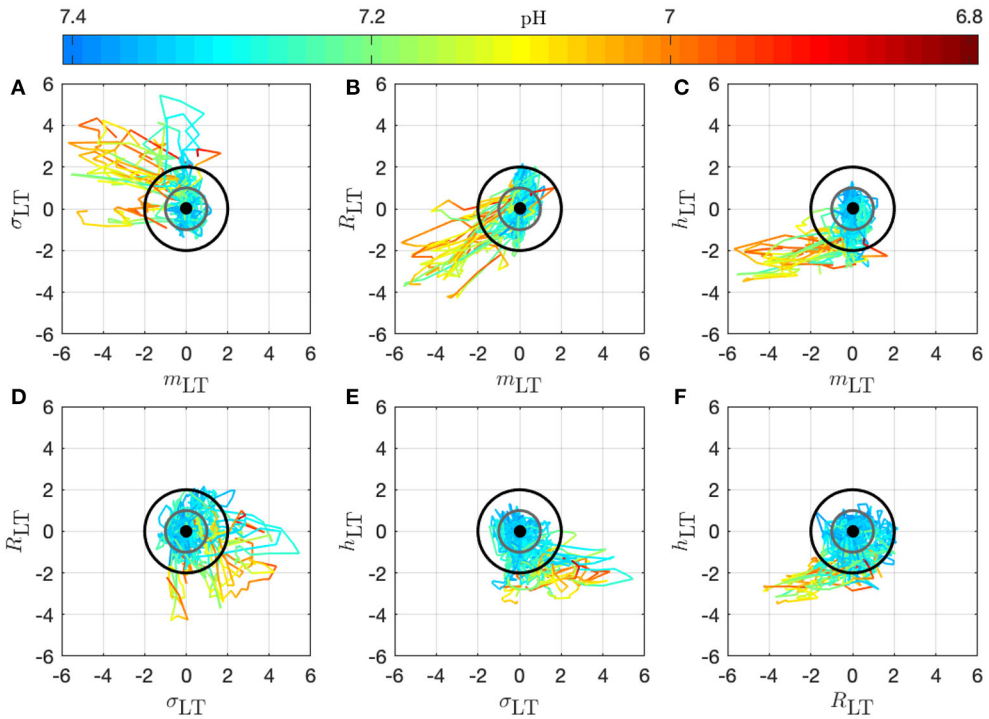

FIGURE 7 | Trajectories of the vector $\vec{u}_k - \vec{u}_0$ for all 14 animals in the phase space; the 6 subplots **(A–F)** correspond to the 6 possible projections onto planes (using 2 coordinates of the vector). Each trajectory corresponds to an animal and is colored to indicate the pH value at the time k: in this way, we observe the joint temporal evolution of $\vec{u}_k$ and of the pH throughout the experiment. Trajectories have been centered by subtracting $\vec{u}_0$, according to Equation (11) to account for the variability between animals: the thick black dot at the origin thus represents the starting point of all trajectories. The gray circle corresponds to $D = 1$ and the black circle to $D = 2$.

The agreement between the CVD time and the distance time is very satisfying:

the difference between t_{CVD} and t_D is not only always smaller than the difference between t_{pH} and t_{CVD}, but also smaller than 20 min, the size of the time-windows we have used.

However, for one animal (number 473360, last line in **Table 4**), a large discrepancy is observed. A closer examination of both the data and our distance for this animal is given in **Figure 10** and allows us to discuss the sensitivity of our measure. We have used the 4 Hz FHR dataset which was also studied in earlier literature. This dataset is obtained from the R-R intervals data at 4 Hz, which is interpolated from the raw ECG-derived R-R intervals data recorded at 1,000 Hz. As can be seen in **Figure 10**, the genuine 1,000 Hz dataset (in red) is missing some values during short intervals and the resampling process, which uses splines interpolation, creates arbitrary values for the 4 Hz FHR dataset (in black) within such intervals. This results in additional values which exhibit large and fast fluctuations which are non-physiological. Whereas, most of these do not impact the value of the distance D (see **Figures 10e,f,g,h**), there is a time interval (at about $t = 172$ s, see **Figures 10b,d** where D is unexpectedly large, reaching a value around 4. This is concomitant with a sharp drop in FHR, as can be seen in **Figures 10a,c**. This sharp drop is exacerbated on the 4 Hz signal compared to the 1,000 Hz signal, and is very localized in time, which leads to a later decrease of D, contrary to the pathological situation reported in **Figures 10g,h** where D remains at a large value. As a consequence, we obtain a false positive sentinel time t_D which corresponds to this event and is hence much earlier than t_{CVD}, although in agreement with previously reported results using the same 4 Hz FHR dataset (19). We conclude that splines interpolation should be avoided, and we suggest instead not to add or create artificial data points when genuine data is not available. Additionally, each of the quantities we propose, and hence the distance D, can still be computed, as they are all robust with respect to missing data, as seen for example in **Figures 8e,f**.

The robustness with respect to missing data is 2-fold. First, all quantities we compute do not require equi-sampled data: this is in contrast to a power spectrum for example, where missing points prevent the estimation and jeopardize the estimated value if using interpolated values. For our quantities, missing data only impacts the number of points used to compute averages, as can be seen in **Figure 2**. Second, having missing data only reduce the number of points over which statistics are computed: a reduced number of points increases the bias and the variance of the estimators. As can be seen in **Figures 10b,d,f,h**, the distance D evolves smoothly which suggest the standard deviation is not strongly impacted. However, one may wonder if an increased bias impacts the reported values, especially when a lot of data is missing. We report in **Table 5** the average fraction of missing data points in a time-window of size T=20 min: increasing the UCO strength is typically associated with an increase of missing

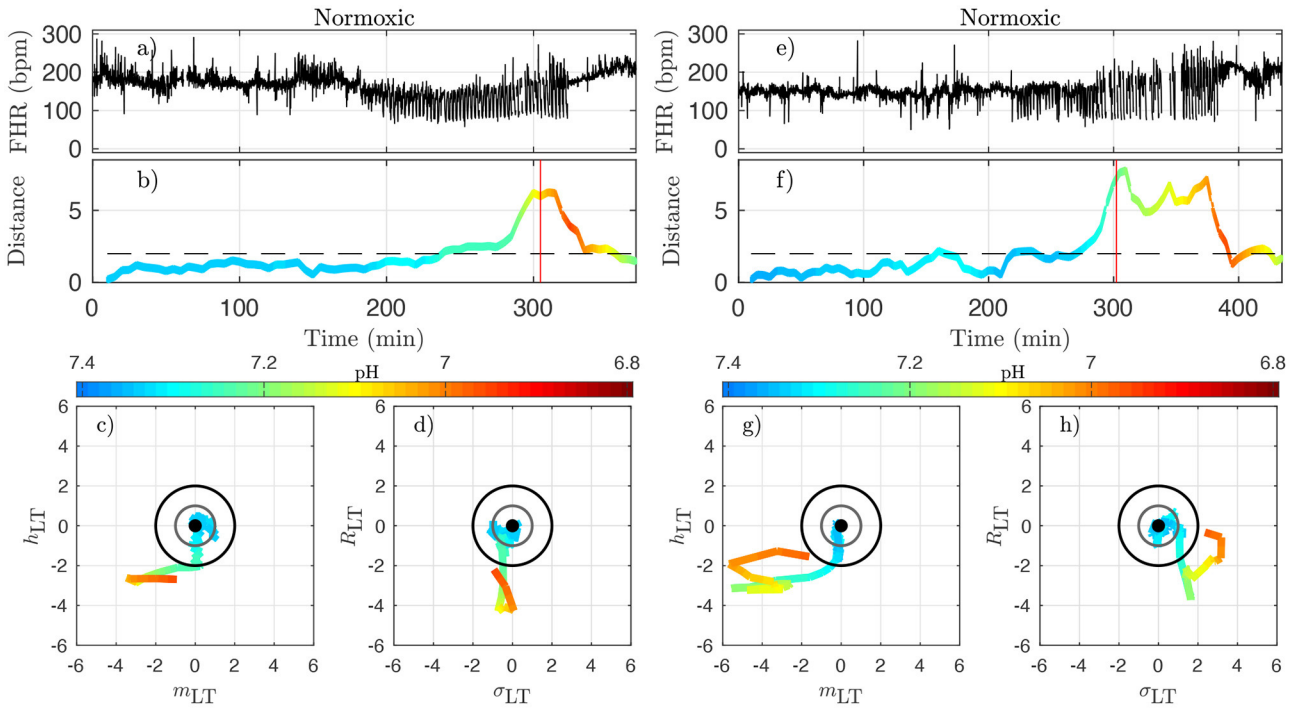

FIGURE 8 | Two examples of normoxic animals (right: animal 473361, left: animal 461060). **(a,e)**: FHR. **(b,f)**: distance D along time; the horizontal dashed line indicates the value $D = 2$ and the red vertical line indicates the CVD time (ABP sentinel), i.e., the time point when we visually confirm the onset of CVD. **(c,g)**: projections of the trajectory of the vector $\vec{u}$ on the plane (m_{LT}, h_{LT}) (the two features that individually better correlate with the pH). **(d,h)**: projections of the trajectory on the plane (σ_{LT}, R_{LT}) (the other two components of $\vec{u}$). The distance D and the trajectories are color-coded with the pH values in each time-window of size T. For clarity, the last part of the trajectory where D decreases below $D = 2$ is omitted.

data. Let's focus on the entropy rate $h_k(\tau)$, which is algorithmicaly the most complex quantity: it has been reported that the bias of $h_k(\tau)$ not only behaves as $1/\sqrt{N}$ (39, 40), similar to the bias of a sliding average over N points like $m_k(\tau)$, but also that this bias is small. A time-window of 20 min should contain $N = 20 \times 60 \times 4 = 4,800$ points, and even a reduction of 50% of available data points should leave more that 2,000 points so a bias smaller than 1% (40). We are thus confident that the reported results are not an indirect measure of the number of missing data points.

A deeper examination of each experiment, using animal's systemic arterial blood pressure data, should clarify the relationship between the increases of the distance from healthy condition and the incipient arterial hypotension. This work is out of scope of the present article which focuses on the dynamics of the clinically relevant 4 Hz-sampled FHR signal.

As such, we propose that the distance D has the potential to serve as an individual biomarker of the incipient CVD, i.e., an early sentinel of the fetal brain injury.

5. CONCLUSIONS AND OUTLOOK

Following achievements in adults and the seminal contribution in Akselrod et al. (41), frequency-based features were used to model linear temporal dynamics in FHR (42–46). To permit richer descriptions of the non-linear dynamics of FHR, information theoretic quantities were used such as entropy rates (34, 47–50), as well as several nonlinear transforms (45, 51–53), and scale-free or (multi)fractal paradigms (54–57). For overviews, interested readers are referred to, e.g., (9, 11, 58–64). An important limitation in the use of these features lies in their dependence on high quality fetal electrocardiogram (ECG) or magnetocardiogram (MCG) data as input. Such data are not readily available in the majority of clinical settings, with over 90% of North American hospitals, for example, still relying on CTG monitors during labor. CTG however provides FHR at a 4 Hz sampling rate, to be compared to 1,000 Hz sampling rate golden standard available with ECG or MCG, while vagally mediated HRV is found on a time scale that goes beyond what is captured at 4 Hz sampling rate. This results in information loss (19, 26, 65, 66). Beyond the mere design of features and their standalone use, numerous efforts were devoted to devise multiple-feature decision rules, often based on supervised learning and machine learning [cf., e.g., (9, 12, 62, 63, 67–69, 69–72)].

In the present work, four FHR features, whose definitions depend on the timescale, are computed on the whole FHR dataset derived from an animal model of human labor to quantify the evolution of FHR temporal dynamics. That means that in our approach we do not rely on considering UCO

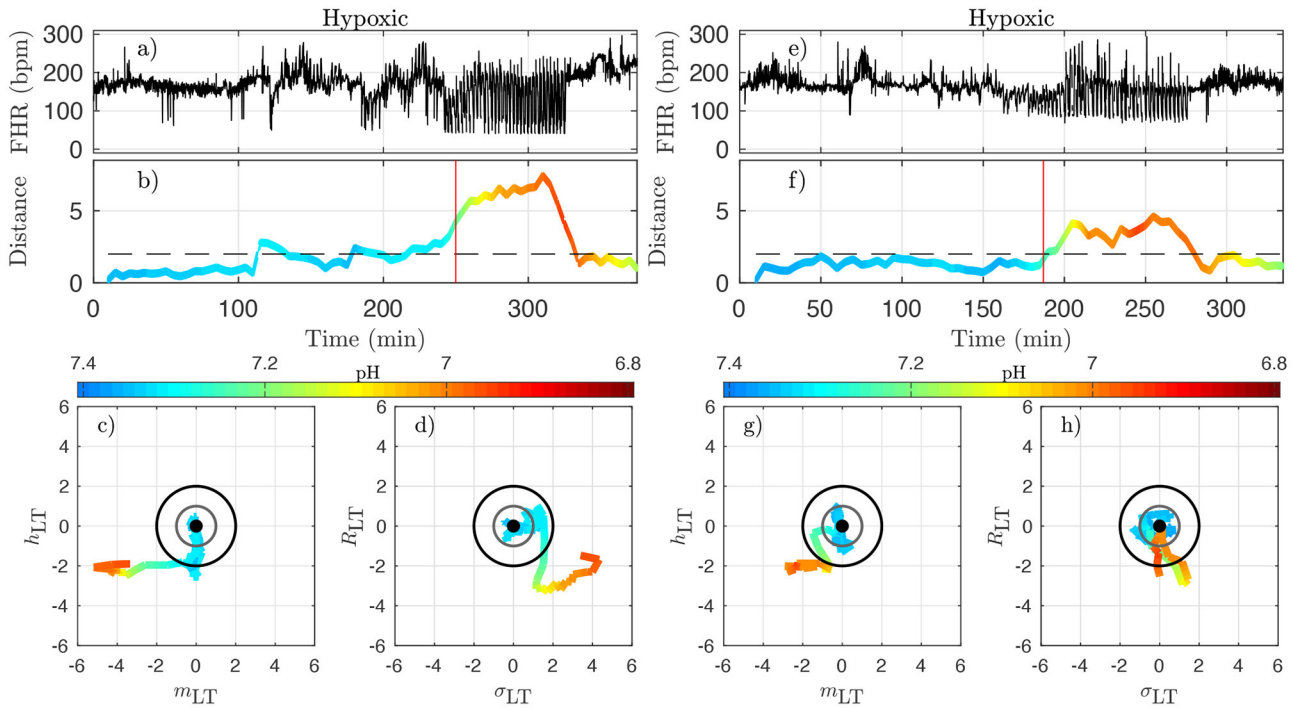

FIGURE 9 | Two examples of hypoxic animals (right: animal 473726, left: animal 473362). See caption of **Figure 8**.

periods only, but are able to process the entire FHR signal as it would be available in real-time in clinical setting. These quantities are local statistical averages that probe the variation, the amplitude of fluctuations and the information content at a given time scale. They are purely statistical quantities that can be computed even when some data points are missing. Firstly, we qualitatively related the variations of such timescale-dependent quantities to the UCO strength; secondly, we quantitatively computed their correlation to metabolite and pH measurements.

As to the etiology of CVD, we propose a role for the Bezold-Jarisch reflex, a vagal cardiac depressor reflex, as part of a complex dynamic interplay, based on the observations of acidemia-triggered inflammation in fetal sheep (16, 17, 25), and studies in adult species linking rising systemic acidemia and inflammation with worsening cardiac contractility, impaired beta-adrenergic and potentiated bradycardic responses (25, 73–77). We suggest that the integrated ability of the four FHR features introduced in this study to track the individual evolution of acidemia and cardiovascular responses stems from capturing the individual complex interplay of the vagally mediated sensing of acidemia and the Bezold-Jarisch reflex, i.e., also vagally mediated intermittent hypotensive ABP responses to UCO-triggered FHR decelerations. This hypothesis needs to be validated in specifically designed animal experiments, for example by repeating the experiments underlying the present study with the variation of performing cervical bilateral, left or right vagotomies. This would allow evaluating the contribution of the vagus nerve to the dynamic interplay between the progressive systemic acidemia, the ensuing systemic inflammatory response, accounting for vagus nerve's lateral asymmetry, to the evolution of FHR decelerations and ABP responses over the period of worsening UCOs comparable in duration to stages 1 and 2 of pushing (18).

We show the relevance of timescales ranging in [2.5 − 8] seconds (equivalently [0.125 − 0.4] Hz in frequencies) for early detection of both acidemia and CVD, matching the scales classically used in FHR analysis and referred to as long-term (57, 63). We observed that reduced pH closely relates to larger m_{LT} which may be interpreted as an increase of baseline FHR (62, 63), and lower entropy rate h_{LT}, in agreement with earlier findings reported in the literature (34). More importantly, a per-individual distance metric was constructed from these four (population-normalized) features to quantify a self-referencing departure from a healthy state for each subject independently. Such a definition raises two issues. Firstly, it requires, as is often the case, that monitoring is started early enough while the fetus is still in a healthy condition, so as to create a self-reference to normal on a per-individual basis. If fetuses are already in distress when monitoring is initiated, the distance, albeit increasing with distress, may fail to detect CVD correctly. Secondly, the definition of the vector, and hence the distance, requires a normalization, which is performed in the present work at the population level, i.e., using an average across subjects. Although such an average should converge rapidly with the population size, this dependence requires further investigations.

It has been documented that sheep fetuses have an individual cardiovascular phenotype in their responses to increasing acidemia due to repetitive intermittent hypoxia (15). Chronically hypoxic fetuses have diminished cardiovascular defenses to hypotensive stress (78). Studying the same dataset, we demonstrated that under the conditions of repetitive UCOs and in comparison to the fetuses who were normoxic on the onset of the UCOs, the hypoxic fetuses exhibit accelerated acidosis (21), altered temporal profile of neuroinflammation following UCOs (20) and deceleration reserve (79). In the

TABLE 4 | Cardiovascular decompensation (CVD) times.

	Animal (ID)	CVD time t_{CVD} (hh:mm)	ΔABP (mmHg)	*D* time t_D (hh:mm)	Delta $t_{CVD} - t_D$ (hh:mm)
Hypoxic	8003	01:55 (03:09)	–6	01:56 (03:10)	–00:01
	473351	01:11 (05:19)	5	01:11 (05:19)	00:00
	473376	01:50 (04:43)	2	02:07 (05:00)	–00:17
	473726	02:01 (04:10)	8	01:55 (04:04)	00:06
	473362	00:59 (03:07)	11	01:16 (03:24)	–00:17
	473352	01:09 (05:08)	1	01:06 (05:05)	00:03
Normoxic	5054	03:25 (04:56)	13	03:29 (05:00)	–00:04
	461060	02:03 (05:02)	-9	01:51 (04:50)	00:12
	5060	02:35 (03:44)	-6	02:20 (03:29)	00:15
	473360	03:41 (05:52)	0	00:44 (02:55)	02:57
	473378	02:07 (05:24)	13	01:58 (05:15)	00:09
	473727	01:34 (03:12)	15	01:51 (03:29)	–00:17
	473377	02:14 (04:42)	6	02:22 (04:50)	–00:08
	473361	03:09 (05:05)	5	02:49 (04:45)	00:20

Comparison of the visually determined vs. computed predictions: t_{CVD} from Gold et al. (27) as reference, and our new distance time t_D, computed by requiring $D > 2.5$ for at least 3 consecutive time-windows, spanning a total duration of 30 min. Times are counted from the first UCO, and values in parenthesis indicate times counted from the beginning of the experiment, to compare with figures. ΔABP indicates the ABP difference at t_{CVD}. The color indicates during which UCO phase CVD occured: mild (green), moderate (magenta) or severe (red). The last column reports the difference $t_{CVD} - t_D$ between the reference CVD time, always earlier than t_{pH}, and the new t_D. Positive values indicate a detection earlier than t_{CVD}. All data are derived from 4 Hz sampled FHR signal.

TABLE 5 | Average fraction of missing data points in the FHR signal in a given regime.

	Animal ID	Baseline	Mild UCO	Moderate UCO	Severe UCO
Hypoxic	8003	1%	2%	6%	22%
	473351	3%	NaN	4%	9%
	473376	2%	1%	3%	8%
	473726	0%	1%	1%	16%
	473362	2%	2%	16%	8%
	473352	5%	NaN	1%	11%
Normoxic	5054	1%	0%	1%	2%
	461060	1%	0%	11%	42%
	5060	9%	20%	1%	23%
	473360	9%	4%	4%	18%
	473378	0%	0%	2%	24%
	473727	5%	3%	8%	17%
	473377	1%	0%	1%	6%
	473361	5%	3%	1%	10%

For each time-window of size $T = 20$ min, we divide the number of missing data points by the expected number of points (=20×60×4), and we then average this ratio over all time-windows available in a given regime.

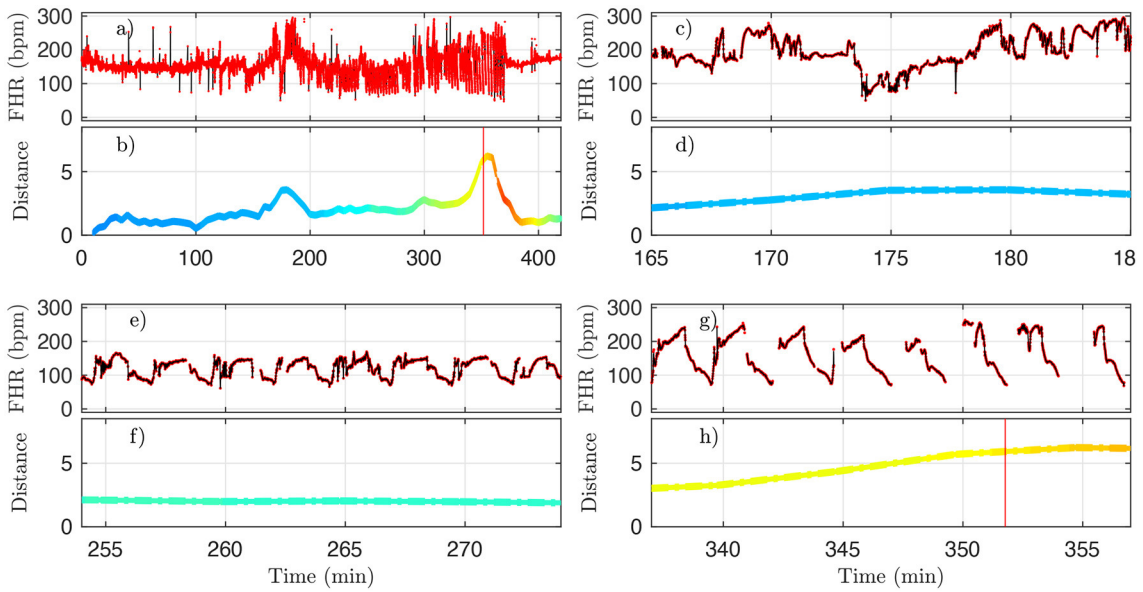

FIGURE 10 | A representative example of FHR and the corresponding distance *D* (colored by pH value) for animal 473360. The 4 Hz dataset used in the analysis is reproduced in black, and the original 1,000 Hz dataset is presented in red. **(a,b)** : complete time trace for all available data. **(c,d)** : zoom in the problematic region, where *D* is unexpectedly large. **(e,f)** : zoom in a region where *D* is small, as expected. **(g,h)** : zoom in a region where *D* is large, as expected. See text.

present study, we could not identify any influence of the phenotype (normoxic/hypoxic) in any of the metrics including the distance D. We are currently exploring whether the initial vector $\vec{u}_0$ may contain such information. Conversely, the finding that the presented approach functions well without the consideration of a pre-existing hypoxia or pattern of labor contractions (UCO severity) is an additional bonus from the clinical viewpoint. Lastly, we recognize that the group of chronically hypoxic animals may have failed to recover from surgical instrumentation adequately, i.e., they were already decompensating rather than becoming "spontaneously" hypoxic for reasons of utero-placental dysfunction preceding the surgery.

Overall, the constructed distance proves able to detect accurately the occurrence of acidemia and CVD from the analysis of FHR only, and without recourse to pH. This opens the route to investigating the relevance of such metrics in clinical practice, as it is non-invasive and much faster than biochemical measurements like pH. Further, for practical purposes, the present studies show that the computation of features and distance is robust to FHR sampled at 4 Hz and, to some extent, to missing data. Also, the FHR features and distance are computed in sliding-windows, permitting an on-line and quasi-real time analysis of the evolution of the dynamics of FHR, and thus in relation to a local health state of the fetus. The extent to which the 20-min sliding-window size, chosen here for proof-of-concept developments, can be further reduced to 10 or 5-min, is under investigation. In conclusion, we propose a real-time FHR-based metric predicting CVD which should be of a great help for health practitioners managing the delivery.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: experimental data may be made available on request (to MF). Requests to access these datasets should be directed to mfrasch@uw.edu.

ETHICS STATEMENT

The animal study was reviewed and approved by University of Western Ontario Council on Animal Care/Canadian Council on Animal Care.

AUTHOR CONTRIBUTIONS

SR designed the signal processing tools, ensured their practical implementation, conducted their application to data, and the analysis and interpretation of the results. He also prepared the figures and tables reported in the article. NBG contributed to the design of the signal processing tools, to the analysis and interpretation of the results, and to the writing of the manuscript. PA contributed to the interdisciplinary connections between signal processing and medical doctor teams, to the interpretation of the results, and to the writing of the manuscript. NG contributed to the measurements and the manuscript. MF designed the experiment, conducted the surgeries, and measurements. He also performed the expert visual detection of the cardiovascular decompensation used as ground truth here. He contributed to writing the article and to the interpretation of the results. All authors contributed to the article and approved the submitted version.

FUNDING

Work supported by Grant ANR-16-CE33-0020 MultiFracs. MF was funded by the Canadian Institutes for Health Research (CIHR grant number 123489).

ACKNOWLEDGMENTS

We thank the Signal Processing and Monitoring Workshop (SPaM workshop) launched under the umbrella of the ANR French Grant #18535 FETUSES where the first stages of the present work were discussed. We gratefully acknowledge Dr. Bryan Richardson and his Perinatal Research Laboratory at the University of Western Ontario for the original design of the animal experiments that enabled the acquisition of the dataset underlying the present study.

REFERENCES

1. Chandraharan E, Arulkumaran S. Prevention of birth asphyxia: responding appropriately to cardiotocograph (CTG) traces. *Best Pract Res Clin Obstet Gynaecol*. (2007) 21:609–24. doi: 10.1016/j.bpobgyn.2007.02.008
2. FIGO. Guidelines for the use of fetal monitoring. *Int J Gynaecol Obstet*. (1986). 25:159–67.
3. Ayres-de Campos D, Spong CY, Chandraharan E, Intrapartum Fetal Monitoring Expert Consensus Panel FIGO. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynaecol Obstet*. (2015) 131:13–24. doi: 10.1016/j.ijgo.2015.06.020
4. Hruban L, Spilka J, Chudek V, Jank P, Huptych M, Bura M, et al. Agreement on intrapartum cardiotocogram recordings between expert obstetricians. *J Eval Clin Pract*. (2015) 21:694–702. doi: 10.1111/jep.12368
5. Frasch M. Saving the brain one heartbeat at a time. *J Physiol*. (2018) 596:5503–4. doi: 10.1113/JP275776
6. Frasch M, Boylan G, Wu H, DD. Commentary: computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Front Physiol*. (2017) 28:721. doi: 10.3389/fphys.2017.00721
7. Gold, N, Frasch, MG. Fetal cerebral perfusion is better than fetal acidaemia for the prediction of brain injury and might be assessable by sophisticated fetal heart rate metrics. *BJOG Int J Obstet Gynaecol*. (2021) 128, 1443. doi: 10.1111/1471-0528.16674
8. 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) 19:289–94. doi: 10.1080/14767050500526172
9. Spilka J, Chudáček V, Koucký M, Lhotská L, Huptych M, Janků P, et al. Using nonlinear features for fetal heart rate classification. *Biomed Signal Proc Control*. (2012) 7:350–7. doi: 10.1016/j.bspc.2011.06.008
10. Nunes I, de Campos DA, Ugwumadu A, Amin P, Banfield P, Nicoll A, et al. Central fetal monitoring with and without computer analysis: a randomized controlled trial. *Obstet Gynecol*. (2017) 129:83–90. doi: 10.1097/AOG.0000000000001799

11. Georgieva A, Abry P, Chudáček V, Djurić PM, Frasch MG, Kok R, et al. Computer-based intrapartum fetal monitoring and beyond: a review of the 2nd workshop on signal processing and monitoring in labor (October (2017). Oxford, UK). *Acta Obstet Gynecol Scand*. (2019) 98:1207–17. doi: 10.1111/aogs.13639
12. Frasch M, Xu Y, Stampalija T, et al. Correlating multidimensional fetal heart rate variability analysis with acid-base balance at birth. *Physiol Meas*. (2014) 35:L1. doi: 10.1088/0967-3334/35/12/L1
13. Cahill AG, Mathur AM, Smyser CD, et al. Neurologic injury in acidemic term infants. *Am J Perinatol*. (2017) 34:668–675. doi: 10.1055/s-0036-1597135
14. Astrup J. Energy-requiring cell functions in the ischemic brain. Their critical supply and possible inhibition in protective therapy. *J Neurosurg*. (1982) 56:482–97. doi: 10.3171/jns.1982.56.4.0482
15. 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:e22100. doi: 10.1371/journal.pone.0022100
16. Frasch MG, Durosier LD, Gold N, Cao M, Matushewski B, Keenliside L, et al. Adaptive shut-down of EEG activity predicts critical acidemia in the near-term ovine fetus. *Physiol Rep*. (2015) 3:129A. doi: 10.14814/phy2.12435
17. Frasch MG, Mansano R, Ross MG, Gagnon R, Richardson BS. Do repetitive umbilical cord occlusions (UCO) with worsening acidemia induce the Bezold-Jarisch reflex (BJR) in the ovine fetus near term? *Reprod Sci*. (2008) 15:129A. doi: 10.1177/1933719108015002102
18. Frasch MG. Letter to the Editor: Bezold Jarisch reflex in the near-term fetus during labor: a matter of time. Am J Physiol Regul Integr Comp Physiol. (2021) 320:R519. doi: 10.1152/ajpregu.00051.2021
19. Gold N, Herry CL, Wang X, Frasch MG. Fetal cardiovascular decompensation during labor predicted from the individual heart rate tracing: a machine learning approach in near-term fetal sheep model. *Front Pediat*. (2021) 9:593889. doi: 10.3389/fped.2021.593889
20. Xu A, Durosier LD, Ross MG, Hammond R, Richardson BS, G FM. Adaptive brain shut-down counteracts neuroinflammation in the near-term ovine fetus. *Front Neurol*. (2014) 5:110. doi: 10.3389/fneur.2014.00110
21. Amaya K, Matushewski B, Durosier L, Frasch M, Richardson B, Ross M. Accelerated acidosis in response to variable fetal heart rate decelerations in chronically hypoxic ovine fetuses. *Am J Obstet Gynecol*. (2016) 2:270. doi: 10.1016/j.ajog.2015.09.084
22. Itskovitz J, LaGamma EF, M RA. Heart rate and blood pressure responses to umbilical cord compression in fetal lambs with special reference to the mechanism of variable deceleration. *Am J Obstet Gynecol*. (1983) 147:451–7. doi: 10.1016/S0002-9378(16)32243-8
23. Richardson BS, Rurak D, Patrick JE, Homan J, Carmichael L. Cerebral oxidative metabolism during sustained hypoxaemia in fetal sheep. *J Dev Physiol*. (1989) 11:37–43.
24. Xu A, Matushewski B, Cao M, Hammond R, Frasch M, Richardson B. The ovine fetal and placental inflammatory response to umbilical cord occlusions with worsening acidosis. *Reprod Sci*. (2015) 22:1409. doi: 10.1177/1933719115580994
25. Prout AP, Frasch MG, Veldhuizen RA, Hammond R, Ross MG, S RB. Systemic and cerebral inflammatory response to umbilical cord occlusions with worsening acidosis in the ovine fetus. *Am J Obstet Gynecol*. (2010) 202:82. doi: 10.1016/j.ajog.2009.08.020
26. Durosier D, Green G, Batkin I, Seely AJ, Ross MG, Richardson BS, et al. Sampling rate of heart rate variability impacts the ability to detect acidemia in ovine fetuses near-term. *Front Pediatr*. (2014) 2:38. doi: 10.3389/fped.2014.00038
27. Gold N, Frasch MG, Herry CL, Richardson BS, Wang X. A doubly stochastic change point detection algorithm for noisy biological signals. *Front Physiol*. (2018) 8:1112. doi: 10.3389/fphys.2017.01112
28. Richman J, Moorman R. Time series analysis using approximate entropy and sample entropy. *Biophys J*. (2000) 78:218A. doi: 10.1152/ajpheart.2000.278.6.H2039
29. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. (2000) 278:H2039–49.
30. Lake DE, Richman JS, Griffin MP, Moorman JR. Sample entropy analysis of neonatal heart rate variability. *Am J Physiol Regul Integr Comp Physiol*. (2002) 283:R789–97. doi: 10.1152/ajpregu.00069.2002
31. Pincus SM. Approximate entropy as a measure of system-complexity. *Proc Natl Acad Sci USA*. (1991) 88:2297–301. doi: 10.1073/pnas.88.6.2297
32. Pincus S. Approximate entropy (ApEn) as a complexity measure. *Chaos*. (1995) 5:110–7. doi: 10.1063/1.166092
33. Spilka J, Roux SG, Garnier NB, Abry P, Goncalves P, Doret M. Nearest-neighbor based wavelet entropy rate measures for intrapartum fetal heart rate variability. In: 2014 36*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Chicago, IL, IEEE (2014). p. 2813–2816.
34. Granero-Belinchon C, Roux SG, Abry P, Doret M, Garnier NB. Information theory to probe intrapartum fetal heart rate dynamics. *Entropy*. (2017) 19:640. doi: 10.3390/e19120640
35. Granero-Belinchon C, Roux SG, Garnier NB, Abry P, Doret M. Mutual information for intrapartum fetal heart rate analysis. *Conf Proc of the IEEE Eng Med Biol Soc*. (2017) 2017:2014–17. doi: 10.1109/EMBC.2017.8037247
36. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. (1948) 27:388–427. doi: 10.1002/j.1538-7305.1948.tb00917.x
37. Frasch MG, Zwiener U, Hoyer D, Eiselt M. Autonomic organization of respirocardial function in healthy human neonates in quiet and active sleep. *Early Hum Dev*. (2007) 83:269–77. doi: 10.1016/j.earlhumdev.2006.05.023
38. David M, Hirsch M, Karin J, Toledo E, Akselrod S. An estimate of fetal autonomic state by time-frequency analysis of fetal heart rate variability. *J Appl Physiol*. (2007) 102:1057–64. doi: 10.1152/japplphysiol.00114.2006
39. Granero-Belinchon C, Roux SG, Abry P, Garnier NB. Probing high-order dependencies with information theory. IEEE Trans Signal Proc. 67:3796–805 (2019). doi: 10.1109/TSP.2019.2920472
40. Granero-Belinchn C, Roux SG, Garnier NB. Information theory for non-stationary processes with stationary increments. Entropy. (2019) 21:1223. doi: 10.3390/e21121223
41. Akselrod S, Gordon D, Ubel FA, Shannon DC, Berger AC, Cohen RJ. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*. (1981) 213:220–2. doi: 10.1126/science.6166045
42. Siira SM, et al. Marked fetal acidosis and specific changes in power spectrum analysis of fetal heart rate variability recorded during the last hour of labour. *BJOG*. (2005) 112:418–23. doi: 10.1111/j.1471-0528.2004.00454.x
43. Gonçalves H, Rocha AP, de Campos DA, Bernardes J. Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery. *Med Biol Eng Comput*. (2006) 44:847–55. doi: 10.1007/s11517-006-0105-6
44. Van Laar JOEH, Porath MM, Peters CHL, Oei SG. Spectral analysis of fetal heart rate variability for fetal surveillance: review of the literature. *Acta Obstet Gynecol Scand*. (2008) 87:300–6. doi: 10.1080/00016340801898950
45. Magenes G, Signorini MG, Ferrario M, Pedrinazzi L, Arduini D. Improving the fetal cardiotocographic monitoring by advanced signal processing. *Conf proc of the IEEE Eng Med Biol Soc*. (2003) 3:2295–8. doi: 10.1109/IEMBS.2003.1280374
46. Siira S, Ojala TH, Vahlberg TJ, Rosn KG, Ekholm EM. Do spectral bands of fetal heart rate variability associate with concomitant fetal scalp pH? *Early Hum Dev*. (2013) 89:739–42. doi: 10.1016/j.earlhumdev.2013.05.007
47. Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett*. (2002) 89:068102. doi: 10.1103/PhysRevLett.89.068102
48. Echeverria JC, Hayes-Gill BR, Crowe JA, Woolfson MS, Croaker GDH. Detrended fluctuation analysis: a suitable method for studying fetal heart rate variability? *Phys Meas*. (2004) 25:763–74. doi: 10.1088/0967-3334/25/3/015
49. Porta A, Bari V, Bassani T, Marchi A, Tassin S, Canesi M, et al. Entropy-based complexity of the cardiovascular control in Parkinson disease: Comparison between binning and k-nearest-neighbor approaches. Conf Proc IEEE Eng Med Biol Soc (2013) 2013:5045–8. doi: 10.1109/EMBC.2013.6610682
50. Spilka J, Roux SG, Garnier NB, Abry P, Goncalves P, Doret M. Nearest-neighbor based wavelet entropy rate measures for intrapartum fetal heart rate variability. In: *Engineering in Medicine and Biology Society (EMBC)*. 2014 36*th Annual International Conference of the IEEE*. Chicago, IL, IEEE (2014). p. 2813–16.

51. Magenes G, Signorini MG, Arduini D. Classification of cardiotocographic records by neural networks. *Proc IEEEINNS-ENNS Int Joint Conf Neural Netw*. (2000) 3:637–41. doi: 10.1109/IJCNN.2000.861394
52. Chudek V, Anden 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) 61:1100–8. doi: 10.1109/TBME.2013.2294324
53. 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) 121:889–94. doi: 10.1111/1471-0528.12568
54. Francis DP, Willson K, Georgiadou P, Wensel R, Davies LC, Coats A, et al. Physiological basis of fractal complexity properties of heart rate variability in man. *J Physiol*. (2002) 542(Pt 2):619–29. doi: 10.1113/jphysiol.2001.013389
55. Doret M, Helgason H, Abry P, Gonçalvés P, Gharib C, Gaucherand P. Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor. *Am J Perinatol*. (2011) 28:259–66. doi: 10.1055/s-0030-1268713
56. Abry P, Roux S, Chudáček V, Borgnat P, Gonçalves P, Doret M. Hurst exponent and intrapartum fetal heart rate: impact of decelerations. In: 26*th International Symposium on Computer-Based Medical Systems (CBMS)*. Porto: IEEE (2013) p. 1–6. doi: 10.1109/CBMS.2013.6627777
57. 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:e0136661. doi: 10.1371/journal.pone.0136661
58. Frasch MG, Mller T, Hoyer D, Weiss C, Schubert H, Schwab M. Nonlinear properties of vagal and sympathetic modulations of heart rate variability in ovine fetus near term. *Am J Physiol Regul Integr Compar Physiol*. (2009) 296:R702–7. doi: 10.1152/ajpregu.90474.2008
59. Frasch MG, Frank B, Last M, Mller T. Time scales of autonomic information flow in near-term fetal sheep. *Front Physiol*. (2012) 3:378. doi: 10.3389/fphys.2012.00378
60. Gierałtowski J, Hoyer D, Tetschke F, Nowack S, Schneider U, Zebrowski J. Development of multiscale complexity and multifractality of fetal heart rate variability. *Auton Neurosci Basic Clin*. (2013) 178:29–36. doi: 10.1016/j.autneu.2013.01.009
61. Haritopoulos M, Illanes A, Nandi A. Survey on Cardiotocography Feature Extraction Algorithms for Fetal Welfare Assessment. Cham: Springer International Publishing (2016).
62. 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*. (2016) 21:664–71. doi: 10.1109/JBHI.2016.2546312
63. Abry P, Spilka J, Leonarduzzi R, Chudáček V, Pustelnik N, Doret M. Sparse learning for intrapartum fetal heart rate analysis. *Biomed Phys Eng Express*. (2018) 4:034002. doi: 10.1088/2057-1976/aabc64
64. Herry CL, Burns P, Desrochers A, Fecteau G, Durosier LD, Cao M, et al. Vagal contributions to fetal heart rate variability: an omics approach. *Physiol Meas*. (2019) 40:065004. doi: 10.1088/1361-6579/ab21ae
65. Li X, Xu Y, Herry C, Durosier LD, Casati D, Stampalija T, et al. Sampling frequency of fetal heart rate impacts the ability to predict pH and BE at birth: a retrospective multi-cohort study. Physiol Meas. (2015) 36:L1–L12. doi: 10.1088/0967-3334/36/5/L1
66. Frasch MG. Heart rate variability code: Does it exist and can we hack it? (2020). https://arxiv.org/abs/2001.08264.
67. Bernardes J, Moura C, de Sa JP, Leite LP. The Porto system for automated cardiotocographic signal analysis. *J Perinat Med*. (1991) 19:61–65. doi: 10.1515/jpme.1991.19.1-2.61
68. Costa A, Ayres-de Campos D, Costa F, Santos C, Bernardes J. Prediction of neonatal acidemia by computer analysis of fetal heart rate and ST event signals. *Am J Obstet Gynecol*. (2009) 201:464.e1–464.e6. doi: 10.1016/j.ajog.2009.04.033
69. 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) 57:771–9. doi: 10.1109/TBME.2009.2035818
70. Georgieva A, Payne SJ, Moulden M, Redman CWG. Artificial neural networks applied to fetal monitoring in labour. *Neural Comput Appl*. (2013) 22:85–93. doi: 10.1007/s00521-011-0743-y
71. Czabanski R, Jezewski J, Matonia A, Jezewski M. Computerized analysis of fetal heart rate signals as the predictor of neonatal acidemia. *Expert Syst Appl*. (2012) 39:11846–60. doi: 10.1016/j.eswa.2012.01.196
72. Xu L, Redman CW, Payne SJ, Georgieva A. Feature selection using genetic algorithms for fetal heart rate analysis. *Phys Meas*. (2014) 35:1357–71. doi: 10.1088/0967-3334/35/7/1357
73. Frasch M, Szynkaruk M, Prout A, Nygard K, Cao M, Veldhuizen R, et al. Decreased neuroinflammation correlates to higher vagus nerve activity fluctuations in near-term ovine fetuses: a case for the afferent cholinergic anti-inflammatory pathway? *J Neuroinflammation*. (2016) 13:103. doi: 10.1186/s12974-016-0567-x
74. Amorim M, de Deus J, Cazuza R, Mota C, da Silva L, Borges G, et al. Neuroinflammation in the NTS is associated with changes in cardiovascular reflexes during systemic inflammation. *J Neuroinflammation*. (2019) 16:125. doi: 10.1186/s12974-019-1512-6
75. Schotola H, Toischer K, Popov A, Renner A, Schmitto J, J G, et al. Mild metabolic acidosis impairs the β-adrenergic response in isolated human failing myoardium. *Crit Care*. (2012) 16:R153. doi: 10.1186/cc11468
76. Mitchell J, Wildenthal K, Johnson RJ. The effects of acid-base disturbances on cardiovascular and pulmonary function. *Kidney Int*. (1972) 1:375–89. doi: 10.1038/ki.1972.48
77. Kimmoun A, Novy E, Auchet T, Ducrocq N, Levy B. Hemodynamic consequences of severe lactic acidosis in shock states: from bench to bedside. *Crit Care*. (2015) 19:175. doi: 10.1186/s13054-017-1624-2
78. Allison BJ, Brain KL, Niu Y, Kane AD, Herrera EA, Thakor AS, et al. Altered cardiovascular defense to hypotensive stress in the chronically hypoxic fetus. *Hypertension*. (2020) 76:1195–207. doi: 10.1161/HYPERTENSIONAHA.120.15384
79. Rivolta M, Stampalija T, Frasch M, Sassi R. Theoretical value of deceleration capacity points to deceleration reserve of fetal heart rate. *IEEE Trans Biomed Eng*. (2020) 67:1176–85. doi: 10.1109/TBME.2019.2932808

Conflict of Interest: MF has an aECG patent WO2018160890A1.

The remaining authors declare that the research was conducted in the absence of any other commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Artificial Intelligence

ORIGINAL RESEARCH
published: 20 August 2021
doi: 10.3389/frai.2021.674238

Estimation and Discriminability of Doppler Ultrasound Fetal Heart Rate Variability Measures

Johann Vargas-Calixto[1]*, *Philip Warrick*[1,2] and *Robert Kearney*[1]

[1]Department of Biomedical Engineering, McGill University, Montreal, QC, Canada, [2]PeriGen Inc., Montreal, QC, Canada

Continuous electronic fetal monitoring and the access to databases of fetal heart rate (FHR) data have sparked the application of machine learning classifiers to identify fetal pathologies. However, most fetal heart rate data are acquired using Doppler ultrasound (DUS). DUS signals use autocorrelation (AC) to estimate the average heartbeat period within a window. In consequence, DUS FHR signals loses high frequency information to an extent that depends on the length of the AC window. We examined the effect of this on the estimation bias and discriminability of frequency domain features: low frequency power (LF: 0.03–0.15 Hz), movement frequency power (MF: 0.15–0.5 Hz), high frequency power (HF: 0.5–1 Hz), the LF/(MF + HF) ratio, and the nonlinear approximate entropy (ApEn) as a function of AC window length and signal to noise ratio. We found that the average discriminability loss across all evaluated AC window lengths and SNRs was 10.99% for LF 14.23% for MF, 13.33% for the HF, 10.39% for the LF/(MF + HF) ratio, and 24.17% for ApEn. This indicates that the frequency domain features are more robust to the AC method and additive noise than the ApEn. This is likely because additive noise increases the irregularity of the signals, which results in an overestimation of ApEn. In conclusion, our study found that the LF features are the most robust to the effects of the AC method and noise. Future studies should investigate the effect of other variables such as signal drop, gestational age, and the length of the analysis window on the estimation of fHRV features and their discriminability.

Keywords: fetal heart rate, cardiotocography, autocorrelation, Doppler ultrasound, classification, fetal heart rate variability

OPEN ACCESS

Edited by:
Martin Gerbert Frasch, University of Washington, United States

Reviewed by:
Faezeh Marzbanrad, Monash University, Australia
Nicolò Pini, Columbia University Irving Medical Center, United States

***Correspondence:**
Johann Vargas-Calixto
carlos.vargascalixto@mail.mcgill.ca

Specialty section:
This article was submitted to Medicine and Public Health, a section of the journal Frontiers in Artificial Intelligence

Received: *01 March 2021*
Accepted: *27 July 2021*
Published: *20 August 2021*

Citation:
Vargas-Calixto J, Warrick P and Kearney R (2021) Estimation and Discriminability of Doppler Ultrasound Fetal Heart Rate Variability Measures. Front. Artif. Intell. 4:674238. doi: 10.3389/frai.2021.674238

INTRODUCTION

Continuous electronic fetal monitoring (EFM) is a standard of care during the antepartum and intrapartum periods (American College of Obstetricians and Gynecologists, 2014). EFM involves measuring two signals: fetal heart rate (FHR) and uterine pressure (UP). These two signals make up what is known as cardiotocography (CTG). Non-invasive Doppler ultrasound (DUS) is the preferred FHR acquisition method in clinical settings (Kupka et al., 2020). Uterine pressure is commonly acquired using external sensors that measure the tension in the maternal abdominal wall (Smyth, 1957). There are other acquisition methods: fetal scalp electrocardiography (ECG) for FHR; and intrauterine probes for uterine pressure (Ayres-De-Campos and Nogueira-Reis, 2016). However, these methods are invasive and are typically used only when external monitoring is not possible.

During the antepartum period, FHR monitoring has been shown to provide information about fetal reactivity (Romano et al., 2006) and abnormalities such as intrauterine growth restriction

(Signorini et al., 2003; Signorini et al., 2020). During labour, the fetus is exposed to repeated periods of hypoxia during uterine contractions (McNamara and Johnson, 1995). If severe enough, sustained hypoxia can lead to metabolic acidosis and hypoxic-ischemic encephalopathy (HIE). Clinicians assess the risk of acidosis and HIE by visually monitoring the EFM for characteristic FHR patterns such as the baseline, accelerations, and decelerations (American College of Obstetricians and Gynecologists, 2014; Lear et al., 2018). Nevertheless, visual assessment of FHR tracings has low specificity and sensitivity as well as high intra- and inter-observer variability (Farquhar et al., 2020). The application of computerized analysis to quantify FHR signals has been proposed to reduce intra- and inter-observer variability (Keith and Greene, 1994). However, recent studies show that automating the analysis of classical FHR patterns does not yield a significant improvement in the detection of acidosis or HIE (Elliott et al., 2010; Clark et al., 2017; Campanile et al., 2018).

It is thought that the development of new FHR indices reflecting the physiological phenomena of acidosis and HIE could improve the ability to identify fetuses at risk (Hamilton and Warrick, 2013). In this context, fetal heart rate variability (fHRV) shows promise to be an important marker of fetal status (Signorini et al., 2003). Heart rate variability (HRV) quantifies variations in the length of the RR interval in successive heartbeats and has been widely used in adults (Acharya et al., 2006). Most HRV analysis algorithms are based on RR intervals derived from ECG signals (Ramshur 2010). However, it is difficult to use these methods for fetal monitoring since DUS measures of FHR do not provide the RR intervals. For this reason, clinical use of fHRV is generally limited to the visual analysis of FHR variations around its baseline.

The DUS transducer emits an ultrasound wave towards the fetal heart. The movement of the fetal heart changes the frequency of the reflected wave due to the Doppler effect (Hamelmann et al., 2020). As a result, both the amplitude and phase of the reflected wave are modulated and consequently its envelope varies with a frequency related to FHR (Hamelmann et al., 2020). FHR is then estimated from the autocorrelation (AC) of the DUS signal envelope computed over a window several seconds long. The AC, which measures the similarity of the signal to itself across time, will have a maximum at a lag equal to the average RR interval (Kupka et al., 2020). FHR is estimated as the inverse of this average RR interval. Fetal monitors use sliding windows to estimate FHR at a uniform sampling rate.

As a result of the averaging associated with computing the AC method, estimates of fHRV features derived from DUS (FHR_{DUS}) will differ from those estimated from RR intervals (FHR_{RRI}). Thus, estimates of power spectral density (PSD) features computed from uniformly sampled HR have been shown to overestimate the low frequency power and underestimate the high frequency power compared to those computed from non-uniformly sampled RR intervals (Clifford and Tarassenko, 2005). Thus, FHR_{DUS} estimates are smoother and have less high frequency (HF) power. Attempts to reconstruct FHR_{RRI} from FHR_{DUS} have not been able to recover the short-term variability features associated with HF fHRV (Cesarelli et al., 2007; Kupka et al., 2020). The errors in fHRV estimates computed for FHR_{DUS} will depend on the AC window length. Longer windows yield more averaging and thus underestimate HF power. Unfortunately, manufacturers of CTG monitors do not disclose the details of their AC algorithms, making it difficult to compare the estimation errors of different monitors.

More sophisticated methods have been proposed to improve the estimation of FHR_{DUS} (Alnuaimi et al., 2017). Peters et al. (2004) used a low-pass filter to roughly estimate the location of the cardiac cycles and defined an AC window that contained only two heart cycles, improving the estimation of spectral features. Similarly, Jezewski et al. (2011) proposed an algorithm which varied AC window length according to an adaptive estimate of beat-to-beat intervals. Valderrama et al. (2019) developed an open-source AC method that optimizes the peak search parameters using Bayesian optimization. Another approach by Katebi et al. (2020) applied unsupervised hidden semi-Markov models to segment the DUS signal for FHR estimation. This approach was able to recover HF features that were very close to those of fECG (Katebi et al., 2020). Despite their improvements, none of these sophisticated methods have yet been applied in bedside monitors (Jezewski et al., 2017).

The availability of large cohorts of perinatal EFM recordings has motivated the development of machine learning (ML) classifiers to improve the early detection of fetal distress and reduce the risk of further injury (Georgieva et al., 2017; Petrozziello et al., 2019). Thus, fHRV features from FHR_{DUS} have been used to identify fetuses with fetal abnormalities using ML and deep learning (DL) (Georgieva et al., 2017; Petrozziello et al., 2019; Signorini et al., 2020). Nevertheless, the discriminability of these algorithms will be adversely affected by errors in the estimation of fHRV features. Durosier et al. (2014) found that the root mean square of the successive differences (RMSSD) of FHR estimated from FHR_{DUS} had worse discriminability than when estimated from FHR_{RRI}. Similarly, it has been suggested that HF FHR_{DUS} features are less discriminative than from FHR_{RRI} (De Jonckheere et al., 2019). The decreased discriminability of fHRV features, along with the undisclosed differences in commercial FHR_{DUS} estimation algorithms, will likely affect the performance of ML classifiers.

This paper analyzes the influence of the AC window length and noise on the estimation and discriminability of some important linear and non-linear fHRV features. These features considered have all been proposed previously for the detection of fetal distress (Signorini et al., 2003). Despite the development of the new sophisticated AC algorithms, we focus on the classical AC method which is the basis of current monitors. The rationale behind this is the desire understand the properties of fHRV computed from EFM data acquired at bedside with current monitors. Thus, our objectives are twofold: 1) To determine how fHRV features computed from FHR_{DUS} differ from those computed from FHR_{RRI}; and 2) To evaluate how these differences influence the ability to classify signals with different fHRV properties. To do so, we explored how different AC window lengths and noise levels affect the estimation of linear PSD features and the nonlinear feature approximate entropy (ApEn). Our results showed that the low frequency power

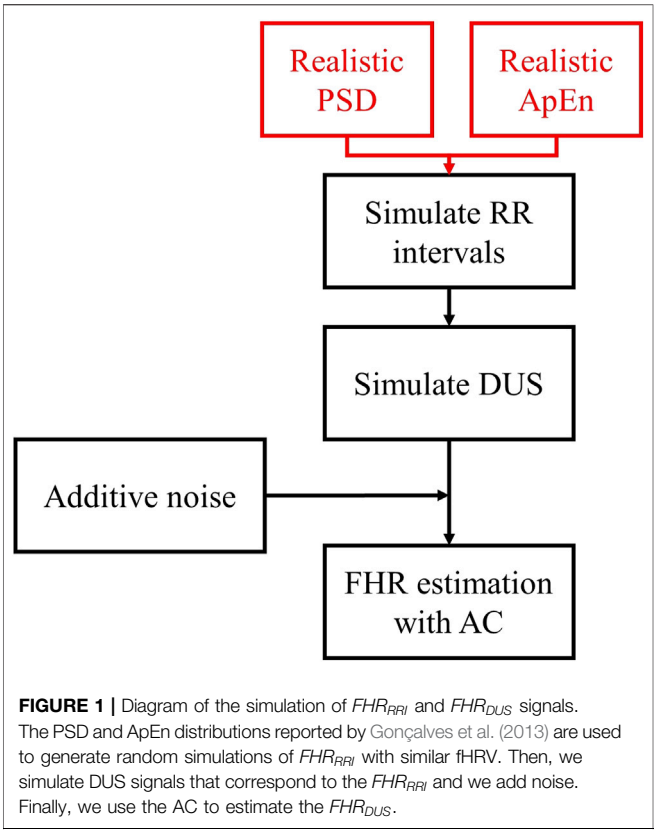

FIGURE 1 | Diagram of the simulation of FHR_{RRI} and FHR_{DUS} signals. The PSD and ApEn distributions reported by Gonçalves et al. (2013) are used to generate random simulations of FHR_{RRI} with similar fHRV. Then, we simulate DUS signals that correspond to the FHR_{RRI} and we add noise. Finally, we use the AC to estimate the FHR_{DUS}.

(LF) is least affected by the AC and noise, while ApEn is affected the most. Furthermore, we examined how the discriminability of each feature varied with AC window length and noise and showed that LF was the most stable feature.

MATERIALS AND METHODS

This section describes the methods used for:

1) Simulating RR intervals and associated DUS signals with PSD and ApEn properties similar to those of normal and acidotic fetuses.
2) Estimating PSD and ApEn features.
3) Evaluating differences between fHRV features estimated from FHR_{DUS} and FHR_{RRI}.
4) Evaluating the discriminability of different simulated fHRV features when applied to normal and acidotic signals.

PhysioNet Fetal ECG Database

We used 80 FHR_{RRI} tracings to validate our simulated RR intervals. These signals were acquired from two databases that included fetal ECG signals and reference annotations indicating the location of the QRS complexes. These annotations were provided by a mixture of experts, volunteers, and specialized algorithms. The first database was acquired by Jezewski et al. and published in PhysioNet (Goldberger et al., 2000; Jezewski et al., 2012). This database contains abdominal and direct fetal ECG records from five term fetuses (gestational ages 38–41 weeks), for 5 minutes each. The second database comprises 75 annotated fetal ECG recordings, each 1 minute long, utilized in the PhysioNet Computing in Cardiology Challenge 2013 (Goldberger et al., 2000; Silva et al., 2013). This database does not indicate the gestational age of the subjects, although the annotations were usually done using simultaneously acquired direct fECG signals. The application of direct fECG is only possible during labor after the rupture of the membranes. The databases do not indicate whether any the fetuses presented any pathological condition. Given the high incidence of normal fetuses, it is likely that the signals were acquired from normal fetuses. The databases also include the location of each R-wave. We used these locations to estimate RR intervals and extracted fHRV features from the RR intervals. We used these fHRV features to validate that our simulations were representative of real data.

Simulation of FHR_{RRI} and FHR_{DUS}

Figure 1 outlines the process for simulating RR intervals, DUS signals, and uniformly sampled FHR. We first generated a sequence of random RR intervals with spectral features for normal or acidosis fetuses similar to those reported by Gonçalves et al., 2013. The 95% confidence intervals (CI) of the power in the low frequency, movement frequency, and high frequency bands reported by Gonçalves et al. are reported in **Table 1**. Afterwards, we generated the DUS envelope signals corresponding to the simulated RR intervals with added noise. Finally, we applied the AC method with a sliding window to generate uniformly sampled FHR_{DUS}.

RR Interval Simulation

We simulated realizations of RR interval sequences, with controlled fHRV PSD structure and nonlinear complexity, as follows:

PSD

We first generated a continuous FHR signal, sampled at 4 Hz, with the desired fHRV spectrum. To do so, we filtered the same white Gaussian noise with three bandpass filters, corresponding to the three bands of interest for fHRV [from (Signorini et al., 2003)]: Low frequency (LF) 0.03–0.15 Hz; Movement frequency (MF) 0.15–0.5 Hz; and High frequency (HF) 0.5–1 Hz band. The three filter outputs were summed in different proportions to generate a signal whose spectrum matched the fHRV spectra reported by Gonçalves et al. (2013).

We then generated a continuous RR interval signal, $RR_C(t)$, from this FHR as $RR = \frac{60}{FHR}$, and upsampled it to 1 kHz using spline interpolation. However, the RR sequence is actually a point process in which the only information of interest is the time of occurrence of an event. Consequently, we transformed the continuous $RR_C(t)$ signal into a point process, $RR_{PP}[i\ ]$, using the method of Clifford and Tarassenko (2005) which proceeds as follows:

1) Sample $RR_C(t)$ at time t_1. Its amplitude, $RR_C(t_1)$, determines the length of the first RR interval. Thus, $RR_{PP}[1\ ] = RR_C(t_1)$. Find the value $RR_C(t_2)$, where $t_2 \geq RR_C(t_1) + t_1$. Then, $RR_{PP}[2\ ] = RR_C(t_2)$.

TABLE 1 | 95% Confidence intervals (CI) for the fHRV estimates reported by Gonçalves et al., 95% CI of the simulated RR intervals fHRV, and the difference of the limits of the 95% CI between the Normal and Acidosis distributions.

	Normal		Acidosis		Difference	
	Gonçalves	Simulated	Gonçalves	Simulated	Gonçalves	Simulated
	95% CI	95% CI	95% CI	95% CI		
LF	19.3	52.78	26.39	129.41	7.09	76.63
	77.21	86.52	264	231.13	186.79	144.61
MF	2.79	2.87	3.36	22.67	0.57	19.8
	13.60	18.96	54.77	110.20	41.17	91.24
HF	0.89	1.63	0.91	10.03	0.02	8.4
	2.25	21.3	8.09	23.35	5.84	2.05
LF/(MF + HF)	4.06	1.80	4.19	1.56	0.13	–0.24
	5.06	9.60	6.19	5.63	1.13	–3.97
ApEn	0.35	0.42	0.25	0.58	–0.1	0.16
	0.52	0.69	0.76	0.79	0.24	0.10

2) Repeat for the length of $RR_C(t)$. At each point $RR_{PP}[i\,] = RR_C(t_i)$ such that $t_i \geq RR_C(t_{i-1}) + t_{i-1}$.

The resulting $RR_{PP}[i\,]$ sequence was fitted to an Autoregressive (AR) model using the Yule-Walker method ("aryule" in the Matlab Signal Processing Toolbox). Multiple RR interval sequences were then generated by filtering independent realizations of white Gaussian noise with this AR model.

Approximate Entropy

ApEn is a measure of signal complexity, and thus random signals will have higher ApEn compared to periodic signals. To control the ApEn of our simulated RR intervals we modified the MIX process of Ferrario et al. (2006). The original MIX process switches randomly between a periodic signal and a uniformly distributed random signal and so does not permit the control of the realization's PSD. To do so, we modified the MIX process to switch randomly between

1) A random sequence $RR_r[i\,]$, with the desired PSD, generated by filtering white noise with the RR AR model.
2) A semi-periodic signal $RR_{sp}[i\,]$ generated by concatenating segments of signal $RR_{r2}[i\,]$ with the desired PSD. Each segment has the same length $l \geq 33$ s and a randomly selected initial point $i_1 \leq l{*}O_f$, where O_f is the overlap factor. This will generate sequences with a limited number of patterns. Varying the overlap makes it possible to generate signals with different values of ApEn but the same PSD.

The MIX process switching is controlled by a binary random variable x, that will have a value of one with probability p, and zero otherwise. Varying p will change ApEn without changing the PSD. The $i$ith RR interval is generated by the MIX process as:

$$RR_{MIX}[i] = x[i]{*}RR_{sp}[i] + (1 - x[i]){*}RR_r[i]$$

DUS Envelope Simulation

Each RR interval sequence was transformed into a corresponding DUS envelope signal, sampled at 1 kHz, as follows:

1) A template DUS envelope cycle DUS_t was selected randomly from 15 available periods of the DUS signal envelope shown in Hamelmann et al. (2020).
2) For each RR interval, the selected DUS_t was stretched or contracted to a length equal to $RR_{MIX}[i\,]$ to give $DUS(t, i)$.
3) Consecutive $DUS(t, i)$ were concatenated to generate the $DUS(t)$ signal.
4) A random additive noise signal, $n(t)$, with a uniform distribution and a LF PSD was generated using an algorithm proposed by Nichols et al. (2010). We limited the power of the noise to 7.7 Hz, the same band of the envelope of the $DUS(t)$ signal.
5) The amplitude of $n(t)$ was varied along each realization to control the signal-to-noise ratio (SNR).
6) Finally, we generated $DUS'(t) = DUS(t) + n(t)$.

Figure 2A shows a segment of a simulated $DUS'(t)$ using the RR intervals from a subject in the PhysioNet Database and a SNR of 20 dB. Separate bursts of activity corresponding to cardiac cycles are apparent.

The AC Method

FHR was estimated from the DUS signal by computing its autocorrelation function (AC). The autocorrelation function of a periodic signal is also periodic with the same period. Consequently, the first non-zero maxima in the AC function will reflect the average RR interval. **Figure 2B** shows the AC coefficient function of the DUS signal in **Figure 2A**, estimated from a 4s window. The first non-zero-lag peak occurs at ~0.5 s indicating an FHR = 120 bpm. Sliding the AC window across the signal with steps of 0.25 s will generate an FHR signal sampled at 4 Hz. The blue curve in **Figure 2C** shows the FHR_{DUS} computed in this way from the signal in **Figure 2A**. (Note that **Figure 2C** covers a longer time span than **Figure 2A**). The black stars show the FHR_{RRI} computed from the original RR intervals for comparison purposes. The AC estimates follow the trend of the FHR_{RRI} but deviate around this trend due to the additive noise.

fHRV Differences Between FHR_{RRI} and FHR_{DUS}

Figure 3 shows the procedure used to compare the fHRV estimates from the RR intervals and DUS FHR.

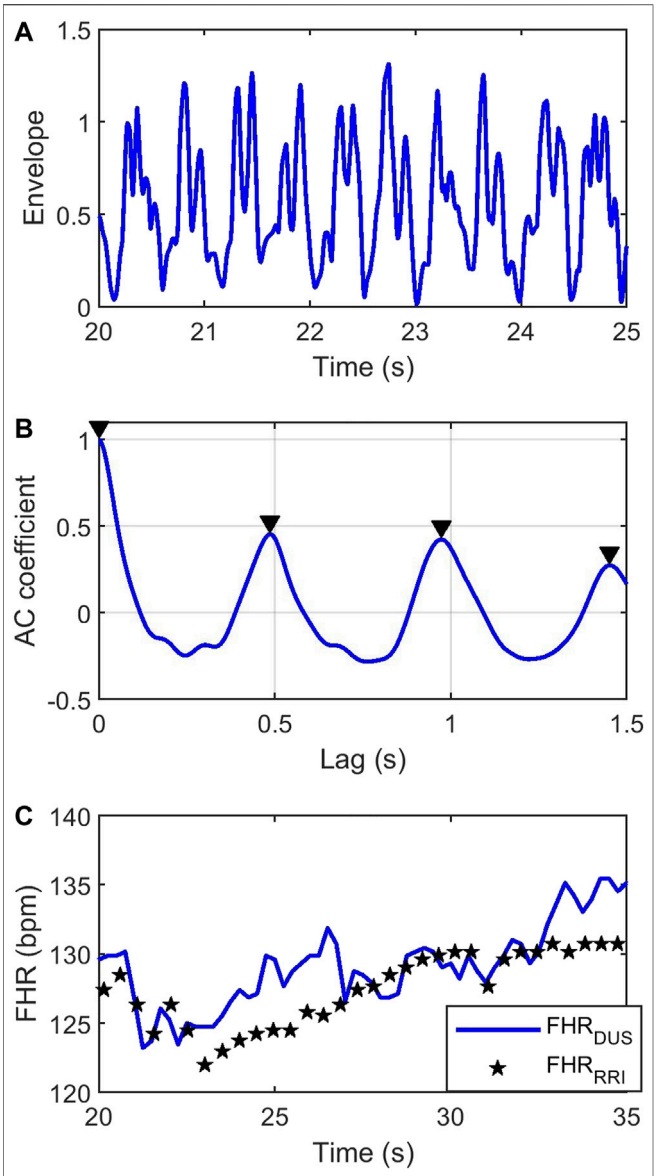

FIGURE 2 | (A) Simulated envelope of the $DUS^{'}(t)$ signal using a series of FHR_{RRI} extracted from the PhysioNet Database and 20 dB SNR. **(B)** AC coefficient (blue) and peaks (black triangles) of the DUS envelope using a 4 s window. **(C)** Simulated FHR_{DUS} (blue), and the non-uniformly sampled FHR_{RRI} (black stars).

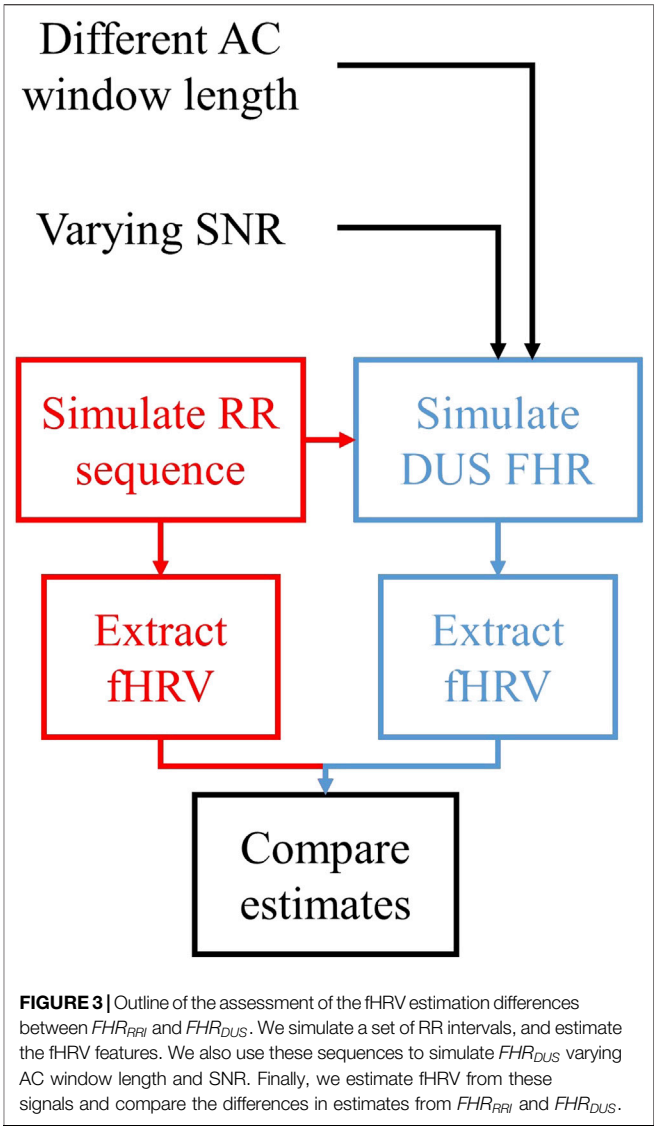

FIGURE 3 | Outline of the assessment of the fHRV estimation differences between FHR_{RRI} and FHR_{DUS}. We simulate a set of RR intervals, and estimate the fHRV features. We also use these sequences to simulate FHR_{DUS} varying AC window length and SNR. Finally, we estimate fHRV from these signals and compare the differences in estimates from FHR_{RRI} and FHR_{DUS}.

1) $RR_{MIX}[i\,]$ sequences were generated with fHRV distributions similar to those reported by Gonçalves et al. for normal and acidotic fetuses.
2) These $RR_{MIX}[i\,]$ sequences were then used to generate corresponding FHR_{DUS} signals and $FHR_{RRI} = \frac{60}{RR_{MIX}}$.
3) fHRV estimates were obtained from FHR_{DUS} and FHR_{RRI}.
4) The estimates were compared as a function of AC window length and SNR.

We simulated 1,000 Monte Carlo (MC) FHR_{RRI} signals having normal and acidotic properties. This yielded a total of 2,000 FHR_{RRI}. For each realization of FHR_{RRI} we generated $DUS^{'}(t)$ signals with 21 SNR values (ranging from −10 to 30 dB in 2 dB steps). These signals were then transformed into FHR_{DUS}, as described above, using 17 AC window lengths (ranging from 1 to 5 in 0.25 s steps). This resulted in 714,000 FHR_{DUS} signals.

FHR Preprocessing

The FHR_{DUS} signals were preprocessed before estimating fHRV features. In some cycles, the additive noise in the DUS signal prevented the peak-finding algorithm from finding the peak that corresponded to the average FHR. To reduce the effect of these outliers, we estimated the moving median of FHR_{DUS} over a 5s window. Estimates that deviated more than 40 bpm from the moving median were removed and replaced by linear interpolation of the adjacent samples. Finally, we limited the estimated FHR_{DUS} to a range of 60–180 bpm.

fHRV Features

The PSDs of the non-uniformly sampled FHR_{RRI}, and the uniformly sampled FHR_{DUS} signals were estimated using the Lomb-Scargle (LS) periodogram as implemented in the function "plomb" in the Matlab Signal Processing Toolbox. We chose the LS periodogram, since it provides unbiased estimates of the power spectrum in non-uniformly sampled signals (Clifford and Tarassenko, 2005). Using alternative methods, such as the Welch periodogram or AR models, would require resampling FHR_{RRI} to a continuous signal, which leads to a biased estimate of the power spectrum (Clifford and Tarassenko, 2005). The normalized power in three frequency bands was then computed as:

$$LF_{pow} = \frac{\sum_{0.03}^{0.15} LS(f)\Delta f}{\sum_{0.03}^{1} LS(f)\Delta f}$$

$$MF_{pow} = \frac{\sum_{0.15}^{0.5} LS(f)\Delta f}{\sum_{0.03}^{1} LS(f)\Delta f}$$

$$HF_{pow} = \frac{\sum_{0.5}^{1} LS(f)\Delta f}{\sum_{0.03}^{1} LS(f)\Delta f}$$

where $LS(f)$ is the PSD estimated using the LS periodogram. In addition, we estimated the LF/(MF + HF) ratio.

ApEn, a measure of the nonlinear complexity of FHR, was estimated as follows:

1) FHR was decimated to 2 Hz, following Gonçalves et al., 2013 who found that sampling the FHR at 2 Hz provided better ApEn estimates than 4 Hz.
2) The function "approximateEntropy" in the Matlab Predictive Maintenance Toolbox was used with an embedding dimension of 2, and radius of 0.2.

Feature Comparison

Differences between features computed from the RR and DUS signals were quantified in terms of their bias and random differences:

$$bd = \frac{E[f_{DUS}] - E[f_{RRI}]}{E[f_{RRI}]} * 100\%$$

$$rd = \frac{\sqrt{\sum (f_{DUS} - f_{RRI})^2}}{E[f_{RRI}]} * 100\%$$

where bd and rd are the normalized bias and random differences, f_{DUS} is a feature estimated from FHR_{DUS}, f_{RRI} is a feature estimated from FHR_{RRI} and $E[x]$ is the expected value.

Discriminability of fHRV

Figure 4 describes the procedure used to assess fHRV discriminability. We simulated normal and acidotic FHR_{RRI}. To remove the effect of the signal amplitude, each realization of FHR_{RRI} was scaled to have a standard deviation of 21.63 bpm, midway between the two reported distributions. Then FHR_{DUS} signals were generated for each FHR_{RRI} realization. PSD features and ApEn were computed for the RR and DUS signals. We constructed a Neyman-Pearson classifier for each signal that used the likelihood ratio of the normal and acidosis distributions, and we estimated the area under the curve (AUC) for the FHR_{RRI} and FHR_{DUS} realizations. The AUC and 95% confidence intervals (CI) were estimated from 1,000 bootstrap samples of the normal and acidotic distributions. To compare features, we computed the following metrics:

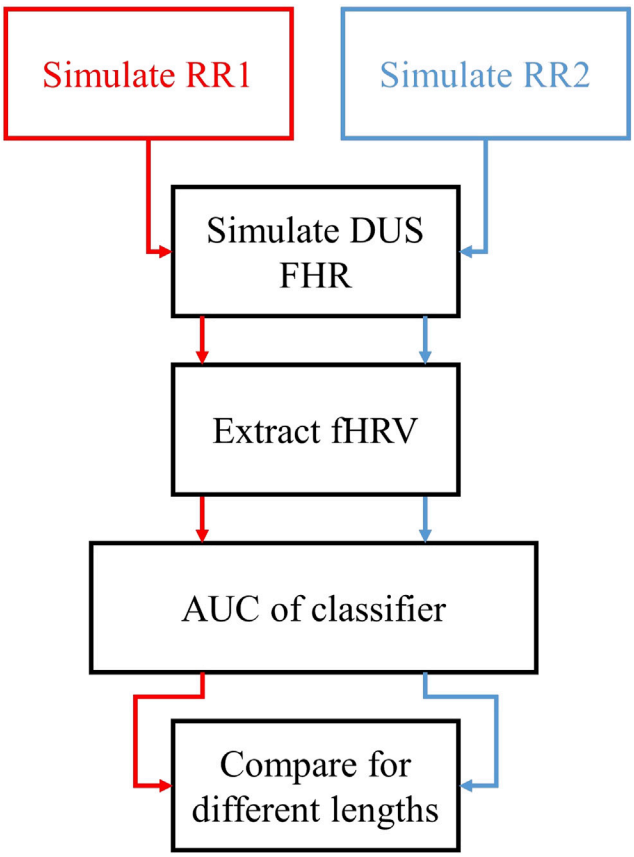

FIGURE 4 | Outline of the assessment of the discriminability of fHRV features as functions of the AC window length. We simulated two sets of FHR_{RRI} sequences with different PSD and ApEn distributions. Then, we extracted the fHRV of the simulated FHR_{DUS} for each case, and we used these estimates to assess the discriminability of each feature using the AUC of Neyman-Pearson classifiers.

1) AUC_{RRI} is the median of the 1000 AUC estimates obtained from bootstrap sampling each pair of fHRV distributions estimated from FHR_{RRI}. $AUC_{DUS}(wl, SNR)$ is the median of the 1000 AUC estimates obtained from bootstrap sampling each pair of fHRV distributions as a function of the window length wl, and the SNR.
2) The normalized difference between AUC_{RRI} and $AUC_{DUS}(wl, SNR)$ given as follows:

$$D_{AUC} = E\left[\frac{AUC_{RRI} - AUC_{DUS}(wl, SNR)}{AUC_{RRI}}\right]$$

where D_{AUC} is the normalized difference, and $AUC_{DUS}(wl, SNR)$ is the AUC_{DUS} as function of AC window length and SNR. D_{AUC}

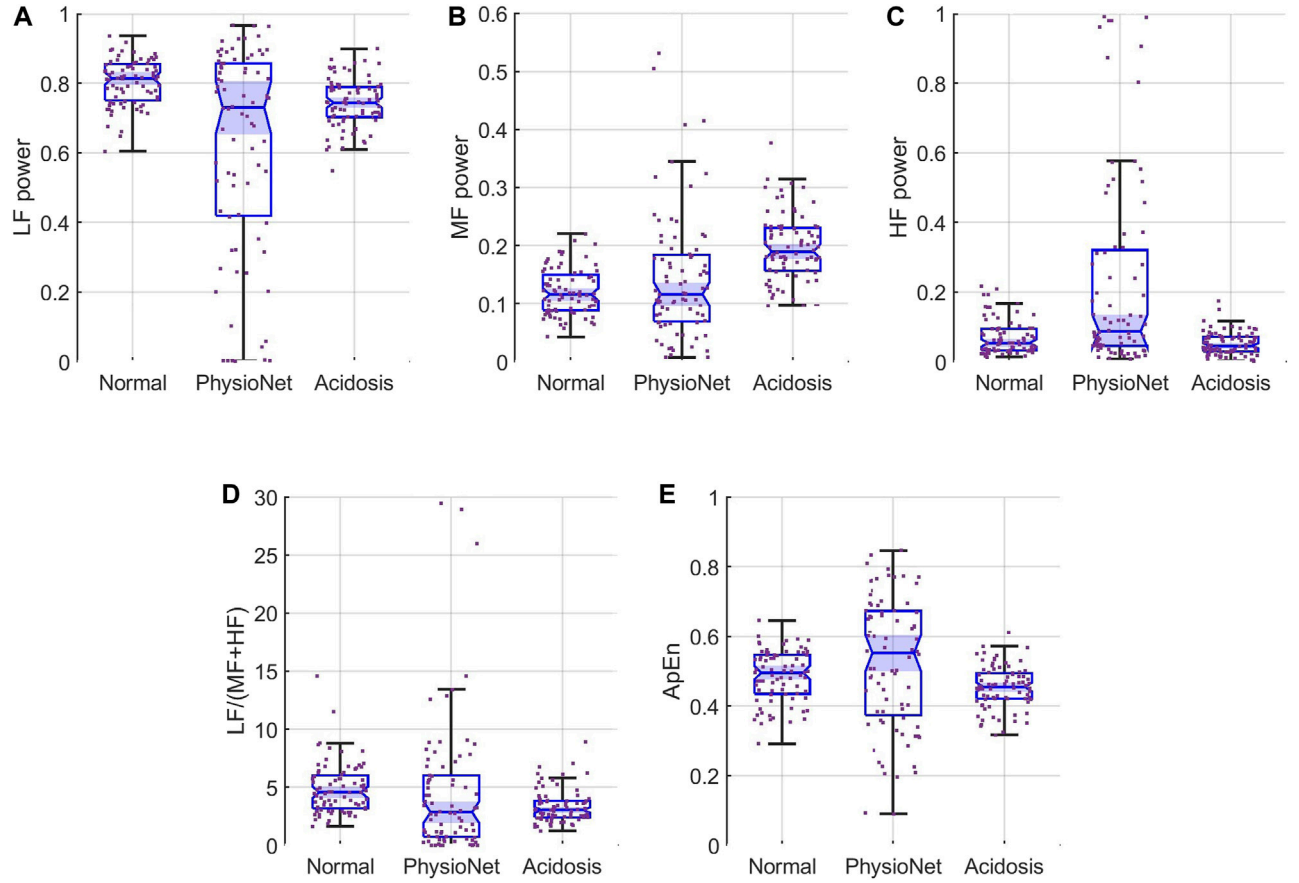

FIGURE 5 | Feature distributions for **(A)** LF, **(B)** MF, and **(C)** HF **(D)** LF/(MF + HF) ratio, and **(E)** ApEn. Each panel show the samples (scattered points) and boxplots for the features for simulated normal (left), the PhysioNet data (middle), and simulated acidosis FHR_{RRI} features (right). The notches, or indentations, in each boxplot indicate the 95% CI for the median of each distribution. These plots show that the PhysioNet data and the simulated normal are not significantly different for any of the features. In contrast, the simulated acidosis distributions are significantly different for the MF and ApEn.

quantifies the mean variation in the discriminability of each feature across all simulated window lengths and SNRs.

3) The normalized standard deviation of AUC_{DUS} for all lengths and SNR given as follows:

$$\sigma_{AUC} = \frac{\sqrt{E\left[\left(AUC_{DUS}(wl, SNR) - E[AUC_{DUS}]\right)^2\right]}}{AUC_{RRI}}$$

where σ_{AUC} is the normalized standard deviation. This metric quantifies how the discriminability of each feature varies as the window length and SNR vary.

RESULTS

Simulation of FHR_{RRI} and FHR_{DUS}

We first compared the features of the simulated FHR_{RRI} sequences to those reported by Gonçalves et al. (2013) for fetuses with normal umbilical cord blood-gas pH (≥7.20) and those with acidotic pH (<7.20). **Table 1** compares the 95% confidence intervals of the PSD and ApEn features reported by Gonçalves et al. to those estimated from our simulated sequences. **Table 1** also shows the difference in the limits of the acidosis and normal distributions. Although the absolute limits of the simulated distributions differ from the reported distributions, they have similar trends; all features except for the LF/(MF + HF) ratio, are larger for the acidosis than the normal class.

We also compared the features of our simulated sequences to those of the 80 subjects in the PhysioNet database. **Figure 5** shows boxplots of the normalized FHR_{RRI} features for three populations: simulated normal (left), PhysioNet data, and simulated acidosis (right). The notches, or indentations, in the box plots indicate the 95% CI of the median of each distribution. From these, it is evident that the medians obtained of the PhysioNet subjects and the simulated normal sequences were not statistically different for any feature. The 95% CI of the medians from the 80 subjects also overlap those of the acidosis FHR_{RRI} except for the MF power and the ApEn. Thus, the distribution of the fHRV features estimated from our simulated FHR_{RRI} were similar to those of real data.

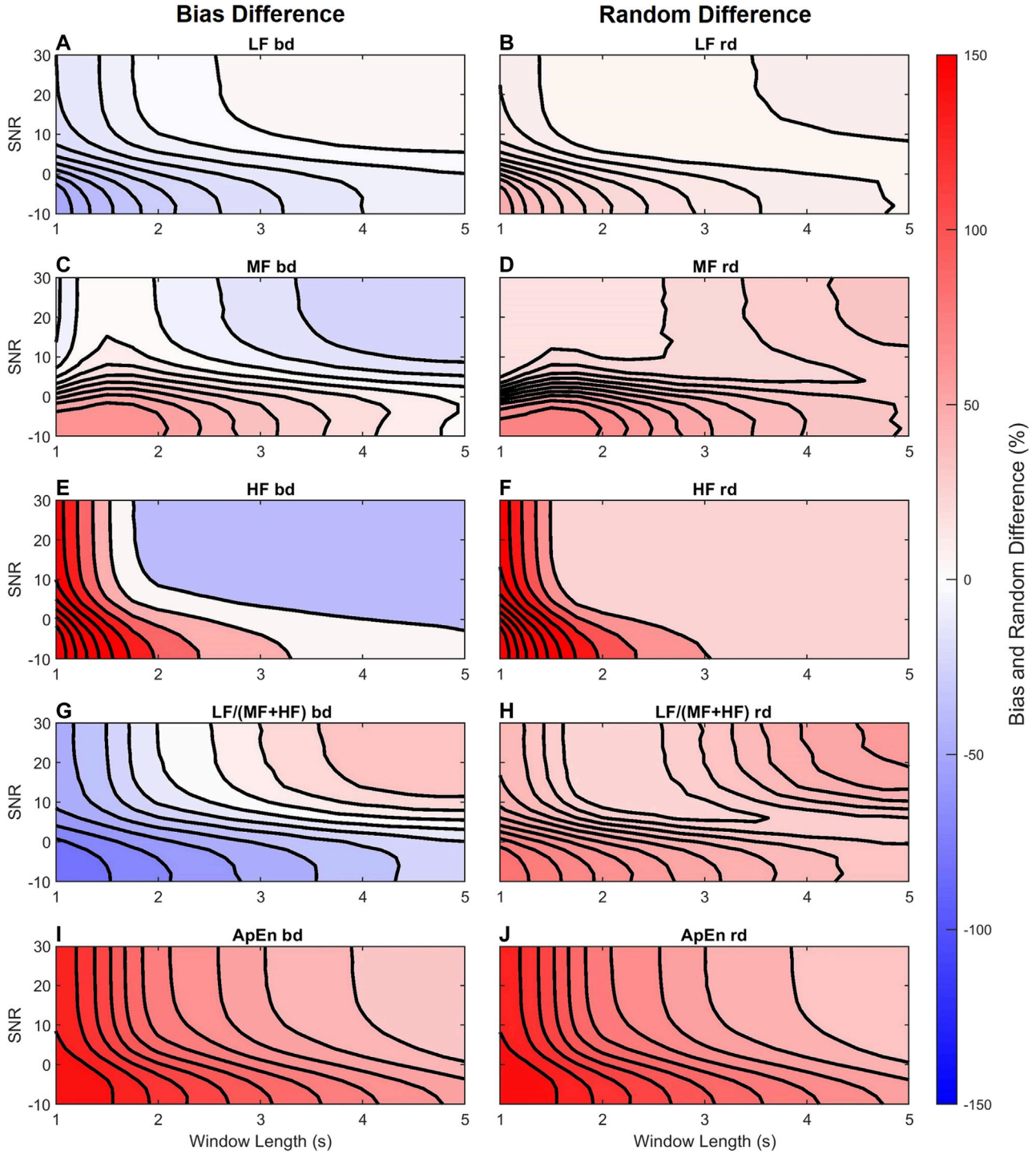

FIGURE 6 | Contour plots of the bias difference, *bd*, and random difference, *rd*, of the LF **(A,B)**, MF **(C,D)**, HF **(E,F)**, LF/(MF + HF) **(G,H)**, and ApEn **(I,J)** features from FHR_{DUS} for the acidosis distributions and varying AC window length (horizontal axis) and SNR (vertical axis). The differences are coded in colors blue (negative), white (zero), and red (positive) according to their magnitude. For visualization, 10 isolines are used in each panel.

TABLE 2 | Median AUC_{RRI} and 95% confidence intervals.

	AUC_{RRI}
LF_{pow}	0.83 (0.81–0.84)
MF_{pow}	0.86 (0.84–0.87)
HF_{pow}	0.82 (0.80–0.83)
$\frac{LF}{MF+HF}$	0.82 (0.80–0.84)
ApEn	0.92 (0.91–0.93)

fHRV Differences Between FHR_{RRI} and FHR_{DUS}

Next we examined the differences between features computed from the simulated FHR_{RRI} and FHR_{DUS}. **Figure 6** shows contour plots for the bias (left column) and random (right column) differences of the five features as functions of window length and SNR. The magnitude of the differences are color coded from blue (negative) to white (zero) and to red (positive). **Figure 6A** shows the LF power is underestimated when the SNR is low or the window length is short; the bias difference is close to zero difference when the window length is longer than 2 s and the SNR is greater than 10 dB. **Figure 6B** shows that the random difference of LF behaved similarly; variability was higher for low SNR and short AC window lengths and decreased as either parameter increased.

Figures 6C,E show that the MF and HF powers were overestimated for low SNR and short AC windows while for long windows and high SNRs they were underestimated. The bias differences had larger magnitude for HF than for MF. Thus, the HF power was most sensitive to the AC window and additive noise. **Figures 6D,F** show that the random difference for both MF and HF were larger for low SNR and short AC windows but decreased as the SNR and window length increased.

Figures 6G,H show that LF/(MF + HF) ratio *bd* and *rd* behaves as expected from the individual trends. Thus, for low SNR and short windows, the ratio was underestimated: smaller LF divided by larger MF and HF estimates produce an underestimated ratio. Similarly, for longer windows and higher SNR, the ratio was overestimated; an almost unbiased LF divided by smaller MF and HF produce an overestimate. The random difference in **Figure 6H** is more complicated to interpret. It was higher for low SNR and short windows, and decreased as either parameter increased. However, it reached a minima at an SNR of 10 dB and window length of 2.5 s and then increased for higher SNRs and window lengths. This might be explained if we consider that the denominator of this ratio (MF + HF) is underestimated in this area. Thus, any variability in the LF estimate, divided by a smaller estimate of (MF + HF) will yield a more variable estimate.

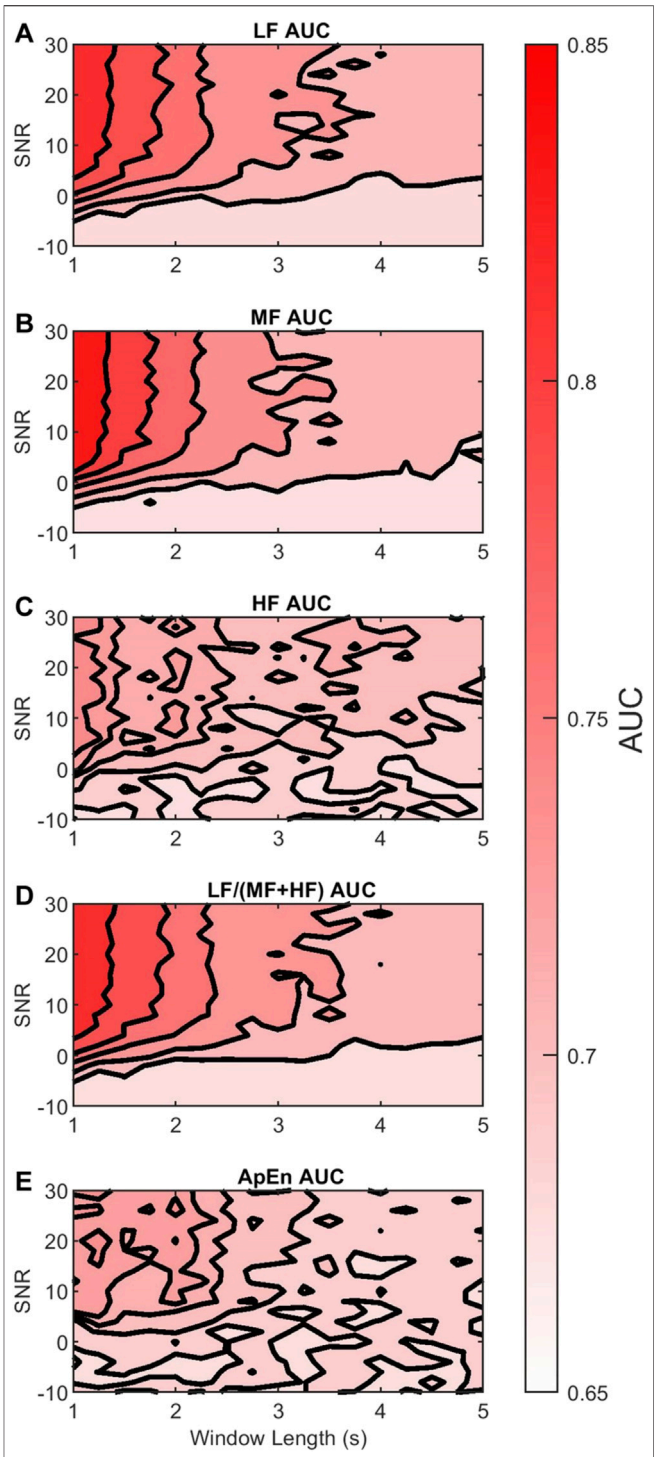

FIGURE 7 | Contour plots of the median AUC_{DUS} of the **(A)** LF, **(B)** MF, **(C)** HF, **(D)** LF/(MF + HF), and **(E)** ApEn features for varying AC window length (horizontal axis) and SNR (vertical axis). The AUC_{DUS} are coded in colors white (0.65), and red (0.85) according to their magnitude. For visualization, 5 isolines are used in each panel.

TABLE 3 | Mean, D_{AUC}, and 95% CI of the difference between AUC_{RRI} and AUC_{DUS} and standard deviation of the estimated AUC_{DUS}, σ_{AUC}, for AC windows 1–5 s long and −10–30 dB SNR.

	Mean difference D_{AUC} (%) and 95% CI	Standard deviation σ_{AUC} (%)
LF_{pow}	−10.99 (−17.29–1.25)	5.02
MF_{pow}	−14.23 (−20.92–−0.40)	5.40
HF_{pow}	−13.33 (−16.39–−8.38)	2.05
$\frac{LF}{MF+HF}$	−10.39 (−16.82–−2.14)	5.15
ApEn	−24.17 (−26.30–−20.18)	1.76

Figure 6I shows that ApEn was always overestimated with the error decreasing as the SNR and window length increased. The random error of ApEn (**Figure 6J**) behaved similarly.

Discriminability of fHRV

We evaluated the discriminability of the features in terms of the AUC of the Neyman-Pearson classifiers for varying AC window length and SNR. As a reference, **Table 2** shows the median and 95% CI AUC_{RRI}. **Figure 7** shows in contour plots the AUC_{DUS} of each of the LF, MF, HF, LF/(MF + HF), and ApEn features. For all cases, AUC_{DUS} decreased with respect to AUC_{RRI} due to the AC method and additive noise. The plots are color coded from the minimum $AUC_{DUS} = 0.65$ (white) to a maximum $AUC_{DUS} = 0.85$ (red). **Figure 7** describes two main trends: 1) for all features, AUC_{DUS} decreases as the SNR decreases, and 2) for all features, AUC_{DUS} decreases as the AC window length increases.

Figures 7A,B show that the discriminability of LF and MF was greatest for SNR larger than 0 dB and windows shorter than 2 s. Thus, in this region, their discriminability was not affected much by the AC method. However, outside this region the color contrast is strong, indicating a large drop in discriminability. In contrast, **Figure 7C** shows that the reduction in HF discriminability was less marked as the window length increased or SNR decreased.

These observations can be contrasted with the results in **Table 3**. **Table 3** shows the 95% CI of the differences between AUC_{RRI} and AUC_{DUS}, the mean difference D_{AUC}, and σ_{AUC} for each feature, AC window 1–5 s long, and −10–30 dB SNR. The D_{AUC} estimates show that in average the LF power loses 10.99%, MF power loses 14.22%, and HF power loses 13.32% of their discriminability. However, the variability of this discriminability, according to σ_{AUC}, is considerably higher for LF and MF (5.02 and 5.40%) than for HF (2.05%). Thus, although HF loses 13.32% of its discriminability due to the AC method and additive noise, the obtained discriminability only varies in 2.05% with respect to the AC window length and the SNR.

Figure 7D shows that the discriminability of LF/(MF + HF) ratio decreases with longer windows and lower SNR. It follows a similar trend to the LF and MF. The estimated D_AUC showed a decrease of 10.39% of its discriminability, and σ_{AUC} showed a variability of 5.15%. These estimates were close to those of the LF power.

Finally, **Figure 7E** shows that the discriminability of ApEn behaved similarly to HF AUC_{DUS}. In this case, D_{AUC} showed the largest loss of discriminability (24.17%) but the smallest σ_{AUC} variability (1.76%). This means that although ApEn loses much of its discriminability due to the AC method, the remaining discriminatory information is affected little by varying SNR or AC window lengths.

DISCUSSION

This paper has two objectives: 1) to analyze differences in fHRV features estimated from FHR_{RRI} and FHR_{DUS}; and 2) to determine how these differences influenced their ability to discriminate between two fHRV distributions. In our analysis, we simulated sequences of RR intervals for which we controlled the PSD and ApEn. Then, we simulated the DUS sampling and AC method, and extracted the relevant features for each objective. Our results indicate that 1) our simulated FHR_{RRI} sequences have fHRV features with distributions similar to those of real data, 2) the estimation of HF power and ApEn are the most affected by the AC method and additive noise, and 3) the loss of discriminability due to the AC method is largest for the ApEn and smallest for the LF power and LF/(MF + HF) ratio. We discuss below each section of these results.

Simulation Issues

The results presented in this paper are based on simulations in which we generated artificial RR intervals and the corresponding DUS signals. The significance of our results will depend on the validity of these simulations. We believe they are valid for the following reasons:

First, an important feature of our simulation of RR intervals was that we were able to generate sequences having both power spectral and entropy features similar to those of real data. **Table 1** and **Figure 5** show that the distribution of fHRV features of our FHR_{RRI} simulations fall within the distributions estimated from the available real data. All features of the simulated RR intervals were comparable to the features estimated from clinical data. This contrasts with previous simulations which controlled only the PSD (Clifford and Tarassenko, 2005), or the entropy independently (Ferrario et al., 2006).

Secondly, we opted to simulate the envelope of the DUS signals rather than the raw DUS signal itself. Raw DUS signals are subject to multiple artifacts during clinical acquisition: movement of the probe or signal loss introduce noise in the signals (Shakespeare et al., 2001; Jezewski et al., 2017). As a solution, fetal monitors use the envelope of the signal, which serves as a LF filter (Hamelmann et al., 2020). This envelope trades the amount of information contained in the signal, such as the location of specific cardiac events (Shakespeare et al., 2001), for robustness in the estimation of the FHR (Hamelmann et al., 2020). Investigating the effect of extracting the envelope of the signal is out of the scope of this study as we focused specifically on the AC method applied to the DUS envelope. Furthermore, we introduced noise in our signal in two ways: 1) we use 15 DUS envelopes reported in the literature (Hamelmann et al., 2020) as templates, which have intrinsic acquisition noise, and 2) we added bandlimited uniform noise, where the cut-off frequency was set to 7.7 Hz. Thus, even for our simulations with the highest SNR, there is noise inherent to the templates that cannot be

removed. This introduces heterogeneity in the signal, each DUS cycle is different from the others.

Finally, it is important to consider the method used to compute FHR_{DUS} from the DUS signal: the estimation of fHRV features will depend on this. We chose to use a standard AC estimation method since this is what is currently used in clinical monitors. Thus, our results are directly relevant to understanding how the properties acquired with current monitors behave. We are aware that a number of more sophisticated methods have been proposed to improve the accuracy of the beat-to-beat estimation of the FHR. These methods are based on the AC method with adaptive parameters, or utilize ML models to extract the beat-to-beat sequence from the DUS signal (Peters et al., 2004; Jezewski et al., 2011; Alnuaimi et al., 2017; Valderrama et al., 2019; Katebi et al., 2020). The effect of those methods on fHRV as function of their parameters is an important question to be explored but is beyond the scope of the present paper.

fHRV Differences Between FHR_{RRI} and FHR_{DUS}

Our experiments aimed to analyze the error of fHRV feature estimation from using the AC method. We found that the length of the autocorrelation window, which determines the extent of signal averaging, had a strong influence on these errors. Longer windows provide more AC averaging, which reduces the effect of additive noise at the cost of beat-to-beat accuracy in the estimation of FHR. In other words, longer averaging windows act as low-pass filters with lower bandwidths. Accordingly, **Figure 6A** shows underestimation of the LF power for short AC windows and low SNR, but the bias difference increases to be almost zero as AC window length and SNR increase. In contrast, the MF and HF powers were increasingly attenuated as the window length increased. As expected, the AC method attenuates the MF and HF power while increasing the relative magnitude of the LF power. This is in agreement with the findings of Clifford and Tarassenko (2005) which showed that interpolated heart rate signals (without averaging) overestimate LF power with respect to higher frequency bands.

Showing a different behavior, ApEn (**Figure 6I**) is always overestimated, which might be due to the effect of additive noise. The ApEn is an estimate of a signal irregularity, and it is higher for random than for periodic signals. Thus, adding random noise increases the signal irregularity which directly increases the ApEn. However, **Figure 6I** shows a decrease when SNR or the window length increase; less noise or more averaging reduces the irregularity in the signal and lowers the ApEn.

In summary, these results show that data from multiple monitors with different parameters may yield different estimates of fHRV. The extent of these differences is documented in our contour plots as a function of window lengths and SNR. Unfortunately, information about the window length used is rarely available for commercial monitors. Unless the manufacturers start to disclose the parameters of their acquisition algorithms, data analysis of such signals must take into account that the variability in the estimated fHRV does not only depend on fetal state but also the CTG monitor.

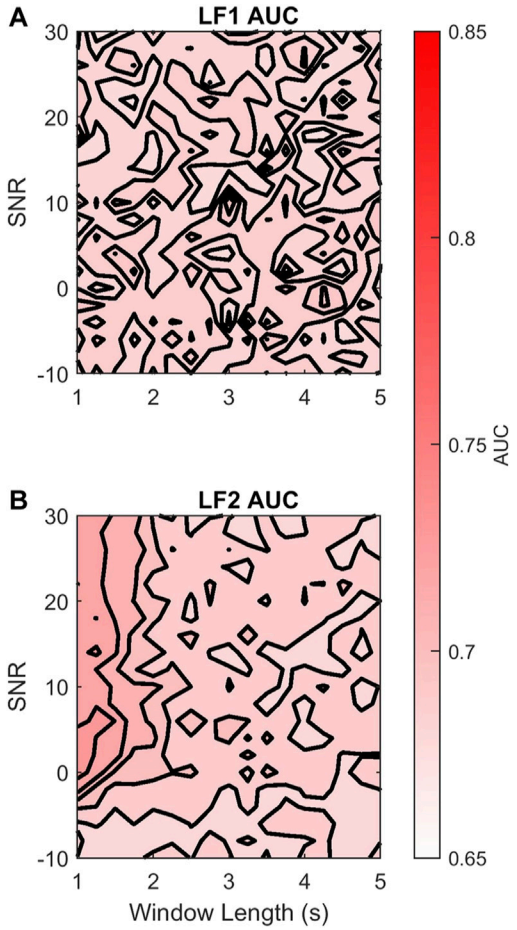

FIGURE 8 | Contour plots of the median AUC_{DUS} of the **(A)** LF1 and **(B)** LF2 sub-bands for varying AC window length (horizontal axis) and SNR (vertical axis). The AUC_{DUS} are coded in colors white (0.65), and red (0.85) according to their magnitude. For visualization, 5 isolines are used in each panel.

Discriminability of fHRV

FHR monitoring during the intrapartum aims to detect fetuses at risk and to use this information to determine whether an emergency cesarean delivery is warranted. Thus, it is important to study how discriminability of certain features is affected by the CTG acquisition methods. Our simulations showed that the discriminability of PSD and ApEn features changed with AC window length and SNR. For all features, the AUC of a Neyman-Pearson classifier decreased as the SNR decreases. This is explained by loss of discriminatory information due to additive noise or large magnitude. Similarly, the AUC decreased as the AC window length increased. This is explained by the loss of discriminatory beat-to-beat information associated with longer AC windows (more averaging).

Two different behaviors can be observed for the five fHRV features analyzed. LF, MF, and LF/(MF + HF) lose less

TABLE 4 | Median AUC_{RRI} and 95% CI, mean difference of AUC_{RRI} and AUC_{DUS}, D_{AUC}, and standard deviation of the estimated AUC_{DUS}, σ_{AUC}, for LF1 and LF2. These estimates were done for AC windows 1 to 5 s long and −10 to 30 dB SNR.

	AUC_{RRI}	Mean difference D_{AUC} (%)	Standard deviation σ_{AUC} (%)
LF1	0.69 (0.67–0.70)	−0.02 (−1.40–1.49)	0.75
LF2	0.72 (0.70–0.74)	−3.38 (−5.48–1.04)	1.74

discriminability on average as defined by D_{AUC}. However, they have higher σ_{AUC} variability across the whole range of AC window lengths and SNRs. This means that under ideal conditions (short AC windows and high SNR) the discriminatory information in these features is well preserved by the AC method. However, samples obtained from monitors that use different AC window lengths could carry quite different discriminatory information, as σ_{AUC} is higher. In contrast, the HF and ApEn features lose more D_{AUC} discriminability on average, but show less σ_{AUC} variability. Thus, these features are more affected by the AC method itself, but they are less dependent on the parameters of this method. This means that when considering data from multiple sources, the LF, MF, and LF/(MF + HF) features might be the best discriminatory parameters if the data is analyzed for each source independently. However, if the data is mixed, then the HF and ApEn features will have the similar discriminatory behavior regardless the source.

The AC method is expected to preserve low frequencies and attenuate high frequencies. Thus, we hypothesized that there might be an LF sub-band that the AC method would not affect. To this end we explored the effects of dividing the LF band into two sub-bands LF1 (0.03–0.072 Hz) and LF2 (0.072–0.15 Hz). **Figure 8A** shows that by doing so, there was no loss in discriminability for LF1 across different AC window lengths and SNR. In contrast, **Figure 8B** shows that the discriminability of LF2 AUC_{DUS} decreased with longer AC windows and lower SNR. **Table 4** quantifies these changes; LF1 loses 0.02% of its D_{AUC} discriminability, and has 0.74% of σ_{AUC} variability across all the range of AC window lengths and SNRs. In contrast, LF2 loses 3.38% of its D_{AUC} discriminability and has 1.74% of σ_{AUC} variability. These results show that the use of the AC method reduces the discriminability of higher frequencies but that frequencies below 72 mHz are not affected by the acquisition method. Therefore, the power in these frequencies provides a discriminatory feature that is independent of the acquisition method and its parameters (AC window length and SNR).

Limitations

We believe these results provide important insight into the effects of computing FHR features using the AC method. Nevertheless there are a number of limitations of the work to consider.

First, the reference distributions that we used to define the normal and acidotic classes were estimated from a handful of cases, which might not be enough. Gonçalves et al. (2013) reported features extracted from 21 normal fetuses and six acidotic fetuses. Thus, a larger database would be necessary to better characterize the distributions of both classes. Furthermore, we use 15 DUS templates in our simulations. Although the use of these templates result in variation of the simulated DUS waves, using a larger number of DUS envelopes as templates might produce more realistic DUS simulations.

Secondly, our model only controls the PSD and ApEn of the simulated FHR_{RRI}. However, it is important to highlight that the power in each band was controlled independently of the others. The fact that the LF/(MF + HF) ratio have a defined distribution suggests that these features are correlated in a way. Our model did not consider this correlation in the features, which resulted in differences between the target distribution and the obtained distribution as shown in **Table 1**. This limitation might have an impact on the interpretation of the results that correspond to the LF/(MF + HF) ratio. In addition, there are many other fHRV features that are used to characterize the variability of FHR signals, namely the short-term variability, long term irregularity, the root mean square of successive differences, among others. It is clear that an ideal simulation would be able to account for all the relevant fHRV features and generate as realistic simulations as possible. Nevertheless, we consider that information available in the PSD of the FHR signal is relatable to some of the time-domain features; the LF power carries information about the long-term evolution of the signal, and the HF power carries information about the short-term beat-to-beat variability of the signal. Similarly, nonlinear indicators of signal irregularity can be related to the ApEn of the signal. Thus, although our study does not control nor account for all the features used in the literature, we consider that our results provide a representative understanding of how different fHRV features behave in response to the AC method and the SNR.

Thirdly, our model does not account for signal loss. Implementing signal loss requires to add a different noise model, which behaves as a switch between signal and no signal. We consider that a future study could expand our model to include such a switch using information from real databases. Parameters such as number of drops, or the duration of the artifact can be characterized in their distributions to generate a realistic DUS signal and FHR_{DUS} estimation.

Fourthly, our model does not account for the nonstationary behavior of intrapartum FHR signals. The simulated FHR signals were time invariant within a window of 10 min. However, real FHR signals are nonlinear and time-varying. For a single subject, it is expected that the fHRV distribution will vary across labor: increased uterine activity will generate responses in the FHR and fHRV (Warrick and Hamilton, 2012; Lear et al., 2018). Thus, the length of analysis window is an important parameter to consider and optimize: short analysis windows will provide large variability in the estimated features, while long analysis windows will include nonstationary FHR. Another factor that will affect the estimated fHRV is fetal state. It has been shown that fetuses have variable fHRV distributions when they are in quiet and active periods (Signorini et al., 2003). Thus, a long analysis window might contain more than one fetal state, which is also nonstationary behavior. An alternative approach to optimizing the length of the analysis window is to consider time-varying or

parameter-varying models to describe FHR and fHRV. Future studies should consider these models in the analysis of intrapartum FHR.

Finally, our model does not consider the effect of gestational age (GA). It has been reported that the distribution of fHRV features vary with GA (Gonçalves et al., 2018). However, no significant difference was reported in PSD or ApEn features for term infants (GA > 36 weeks). Considering that our simulations took as reference fHRV distributions from term infants or intrapartum signals, we believe that our results are valid regardless GA in term infants. Further studies should analyze if the trends of the *bd* and *rd* are different when GA < 36 weeks.

CONCLUSION

Our results demonstrate the susceptibility of fHRV features to the AC method and additive noise in the clinical acquisition of FHR. The dependency of the estimation error on the AC window length, which is part of the proprietary information of the FHR monitor manufacturers, is a limitation in comparing data acquired from different monitors. There is an increasing interest in applying machine learning techniques to FHR tracings on large databases to identify fetuses at risk during antepartum (Signorini et al., 2020) and intrapartum monitoring (Georgieva et al., 2017; Petrozziello et al., 2019). Although the discriminability of fHRV features depends on the AC window length of the FHR monitor and the SNR, it has low variability (<5.4%). Moreover, a feature based on the power below 72 mHz is not affected by the AC method. Thus, understanding the effects of the AC method on fHRV discriminability would potentially lead to a better implementation of ML classifiers of FHR signals when dealing with multiple sources. LF power, MF power, and the LF/(MF + HF) ratio are least affected by the AC method in average but are more influenced by changes in the AC window length and SNR. Classifiers based on these features would benefit from including the fetal monitor model, or acquisition center, as part of the regressor. On the other hand, HF power and ApEn experience the largest loss of discriminability in average, but with lower dependency on AC window length and SNR. Thus, classifiers based on these features would not need to account for differences in the acquisition fetal monitors.

DATA AVAILABILITY STATEMENT

The fetal ECG signals used in this article are openly available as part of the Noninvasive Fetal ECG - The PhysioNet Computing in Cardiology Challenge 2013 (Accession Number: 25401167, URL: https://physionet.org/content/challenge-2013/1.0.0/) and the Abdominal and Direct Fetal ECG Database (Accession Number: 25854665, URL: https://physionet.org/content/adfecgdb/1.0.0/).

AUTHOR CONTRIBUTIONS

JV-C is the first author and the research is part of his PhD. His supervisors PW and RK supported the process during weekly discussions and actively participated in writing the manuscript along with JV-C.

FUNDING

This study was carried thanks to the funding provided by the Bill & Melinda Gates Foundation and the National Institutes of Health.

REFERENCES

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., and Suri, J. S. (2006). Heart Rate Variability: a Review. *Med. Biol. Eng. Comput.* 44, 1031–1051. doi:10.1007/s11517-006-0119

Alnuaimi, S. A., Jimaa, S., and Khandoker, A. H. (2017). Fetal Cardiac Doppler Signal Processing Techniques: Challenges and Future Research Directions. *Front. Bioeng. Biotechnol.* 5, 82. doi:10.3389/fbioe.2017.00082

American College of Obstetricians and Gynecologists (2014). "Intrapartum Considerations and Assessment," in *Neonatal Encephalopathy and Neurologic Outcome, Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Encephalopathy* (Washington, DC: American Academy of Pediatrics), 87–114.

Ayres-De-Campos, D., and Nogueira-Reis, Z. (2016). Technical Characteristics of Current Cardiotocographic Monitors. *Best Pract. Res. Clin. Obstet. Gynaecol.* 30, 22–32. doi:10.1016/j.bpobgyn.2015.05.005

Campanile, M., D'Alessandro, P., Della Corte, L., Saccone, G., Tagliaferri, S., Arduino, B., et al. (2018). Intrapartum Cardiotocography with and without Computer Analysis: a Systematic Review and Meta-Analysis of Randomized Controlled Trials. *J. Maternal-Fetal Neonatal Med.* 33, 2284–2290. doi:10.1080/14767058.2018.1542676

Cesarelli, M., Romano, M., Bifulco, P., Fedele, F., and Bracale, M. (2007). An Algorithm for the Recovery of Fetal Heart Rate Series from CTG Data. *Comput. Biol. Med.* 37, 663–669. doi:10.1016/j.compbiomed.2006.06.003

Clark, S. L., Hamilton, E. F., Garite, T. J., Timmins, A., Warrick, P. A., and Smith, S. (2017). The Limits of Electronic Fetal Heart Rate Monitoring in the Prevention of Neonatal Metabolic Acidemia. *Am. J. Obstet. Gynecol.* 216. doi:10.1016/j.ajog.2016.10.009

Clifford, G. D., and Tarassenko, L. (2005). Quantifying Errors in Spectral Estimates of HRV Due to Beat Replacement and Resampling. *IEEE Trans. Biomed. Eng.* 52, 630–638. doi:10.1109/tbme.2005.844028

Durosier, L. D., Green, G., Batkin, I., Seely, A. J., Ross, M. G., Richardson, B. S., et al. (2014). Sampling Rate of Heart Rate Variability Impacts the Ability to Detect Acidemia in Ovine Fetuses Near-Term. *Front. Pediatr.* 2. doi:10.3389/fped.2014.00038

Elliott, C., Warrick, P. A., Graham, E., and Hamilton, E. F. (2010). Graded Classification of Fetal Heart Rate Tracings: Association with Neonatal Metabolic Acidosis and Neurologic Morbidity. *Am. J. Obstet. Gynecol.* 202, 251–258. doi:10.1016/j.ajog.2009.06.026

Farquhar, C. M., Armstrong, S., Masson, V., Thompson, J. M. D., and Sadler, L. (2020). Clinician Identification of Birth Asphyxia Using Intrapartum Cardiotocography Among Neonates with and without Encephalopathy in New Zealand. *JAMA Netw. Open* 3, e1921363. doi:10.1001/jamanetworkopen.2019.21363

Ferrario, M., Signorini, M. G., Magenes, G., and Cerutti, S. (2006). Comparison of Entropy-Based Regularity Estimators: Application to the Fetal Heart Rate Signal for the Identification of Fetal Distress. *IEEE Trans. Biomed. Eng.* 53, 119–125. doi:10.1109/tbme.2005.859809

Georgieva, A., Redman, C. W. G., and Papageorghiou, A. T. (2017). Computerized Data-Driven Interpretation of the Intrapartum Cardiotocogram: a Cohort Study. *Acta Obstet. Gynecol. Scand.* 96, 883–891. doi:10.1111/aogs.13136

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet. *Circulation* 101, e215-e220. doi:10.1161/01.cir.101.23.e215

Gonçalves, H., Amorim-Costa, C., Ayres-De-Campos, D., and Bernardes, J. (2018). Evolution of Linear and Nonlinear Fetal Heart Rate Indices throughout Pregnancy in Appropriate, Small for Gestational Age and Preterm Fetuses: A Cohort Study. *Comput. Methods Programs Biomed.* 153, 191–199. doi:10.1016/j.cmpb.2017.10.015

Gonçalves, H., Costa, A., Ayres-De-Campos, D., Costa-Santos, C., Rocha, A. P., and Bernardes, J. (2013). Comparison of Real Beat-To-Beat Signals with Commercially Available 4 Hz Sampling on the Evaluation of Foetal Heart Rate Variability. *Med. Biol. Eng. Comput.* 51, 665–676. doi:10.1007/s11517-013-1036-7

Hamelmann, P., Vullings, R., Kolen, A. F., Bergmans, J. W. M., Van Laar, J. O. E. H., Tortoli, P., et al. (2020). Doppler Ultrasound Technology for Fetal Heart Rate Monitoring: A Review. *IEEE Trans. Ultrason. Ferroelect., Freq. Contr.* 67, 226–238. doi:10.1109/tuffc.2019.2943626

Hamilton, E. F., and Warrick, P. A. (2013). New Perspectives in Electronic Fetal Surveillance. *J. Perinat Med.* 41, 83–92. doi:10.1515/jpm-2012-0024

Jezewski, J., Matonia, A., Kupka, T., Roj, D., and Czabanski, R. (2012). Determination of Fetal Heart Rate from Abdominal Signals: Evaluation of Beat-To-Beat Accuracy in Relation to the Direct Fetal Electrocardiogram. *Biomedizinische Technik/Biomedical Eng.* 57, 383–394. doi:10.1515/bmt-2011-0130

Jezewski, J., Roj, D., Wrobel, J., and Horoba, K. (2011). A Novel Technique for Fetal Heart Rate Estimation from Doppler Ultrasound Signal. *BioMedical Eng. OnLine* 10, 92. doi:10.1186/1475-925x-10-92

Jezewski, J., Wrobel, J., Matonia, A., Horoba, K., Martinek, R., Kupka, T., et al. (2017). Is Abdominal Fetal Electrocardiography an Alternative to Doppler Ultrasound for FHR Variability Evaluation? *Front. Physiol.* 8, 305. doi:10.3389/fphys.2017.00305

Jonckheere, J., Garabedian, C., Charlier, P., Storme, L., Debarge, V., and Logier, R. (2019). Influence of Averaged Fetal Heart Rate in Heart Rate Variability Analysis. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2019, 5979–5982. doi:10.1109/EMBC.2019.8856803

Katebi, N., Marzbanrad, F., Stroux, L., Valderrama, C. E., and Clifford, G. D. (2020). Unsupervised Hidden Semi-markov Model for Automatic Beat Onset Detection in 1D Doppler Ultrasound. *Physiol. Meas.* 41, 085007. doi:10.1088/1361-6579/aba006

Keith, R. D. F., and Greene, K. R. (1994). 4 Development, Evaluation and Validation of an Intelligent System for the Management of Labour. *Baillière's Clin. Obstet. Gynaecol.* 8, 583–605. doi:10.1016/s0950-3552(05)80200-7

Kupka, T., Matonia, A., Jezewski, M., Horoba, K., Wrobel, J., and Jezewski, J. (2020). Coping with Limitations of Fetal Monitoring Instrumentation to Improve Heart Rhythm Variability Assessment. *Biocybernetics Biomed. Eng.* 40, 388–403. doi:10.1016/j.bbe.2019.12.005

Lear, C. A., Wassink, G., Westgate, J. A., Nijhuis, J. G., Ugwumadu, A., Galinsky, R., et al. (2018). The Peripheral Chemoreflex: Indefatigable Guardian of Fetal Physiological Adaptation to Labour. *J. Physiol.* 596, 5611–5623. doi:10.1113/jp274937

McNamara, H., and Johnson, N. (1995). The Effect of Uterine Contractions on Fetal Oxygen Saturation. *BJOG:An Int. J. O&G* 102, 644–647. doi:10.1111/j.1471-0528.1995.tb11403.x

Nichols, J. M., Olson, C. C., Michalowicz, J. V., and Bucholtz, F. (2010). A Simple Algorithm for Generating Spectrally Colored, Non-gaussian Signals. *Probabilistic Eng. Mech.* 25, 315–322. doi:10.1016/j.probengmech.2010.01.005

Peters, C. H. L., Broeke, E. D. M. t., Andriessen, P., Vermeulen, B., Berendsen, R. C. M., Wijn, P. F. F., et al. (2004). Beat-to-beat Detection of Fetal Heart Rate: Doppler Ultrasound Cardiotocography Compared to Direct ECG Cardiotocography in Time and Frequency Domain. *Physiol. Meas.* 25, 585–593. doi:10.1088/0967-3334/25/2/015

Petrozziello, A., Redman, C. W. G., Papageorghiou, A. T., Jordanov, I., and Georgieva, A. (2019). Multimodal Convolutional Neural Networks to Detect Fetal Compromise during Labor and Delivery. *IEEE Access* 7, 112026–112036. doi:10.1109/access.2019.2933368

Ramshur, J. T. (2010). *Design, Evaluation, and Application of Heart Rate Variability Analysis Software (HRVAS).* [Master's thesis]. Memphis, TN: University of Memphis.

Romano, M., Bifulco, P., Cesarelli, M., Sansone, M., and Bracale, M. (2006). Foetal Heart Rate Power Spectrum Response to Uterine Contraction. *Med. Bio Eng. Comput.* 44, 188–201. doi:10.1007/s11517-006-0022-8

Signorini, M. G., Magenes, G., Cerutti, S., and Arduini, D. (2003). Linear and Nonlinear Parameters for the Analysis of Fetal Heart Rate Signal from Cardiotocographic Recordings. *IEEE Trans. Biomed. Eng.* 50, 365–374. doi:10.1109/tbme.2003.808824

Signorini, M. G., Pini, N., Malovini, A., Bellazzi, R., and Magenes, G. (2020). Integrating Machine Learning Techniques and Physiology Based Heart Rate Features for Antepartum Fetal Monitoring. *Comput. Methods Programs Biomed.* 185, 105015. doi:10.1016/j.cmpb.2019.105015

Silva, I., Behar, J., Sameni, R., Zhu, T., Oster, J., Clifford, G. D., et al. (2013). Noninvasive Fetal ECG: The PhysioNet/Computing in Cardiology Challenge 2013. *Comput. Cardio.* 40, 149–152.

Shakespeare, S. A., Crowe, J., Hayes-Gill, B. R., Bhogal, K., and James, D. K. (2001). The information content of Doppler ultrasound signals from the fetal heart. *Med. Biol. Eng. Comput.* 39, 619–626. doi:10.1007/BF02345432

Smyth, C. N. (1957). The Guard-Ring Tocodynamometer. Absolute Measurement of Intra-amniotic Pressure by a New Instrument. *BJOG:An Int. J. O&G* 64, 59–66. doi:10.1111/j.1471-0528.1957.tb02599.x

Valderrama, C. E., Stroux, L., Katebi, N., Paljug, E., Hall-Clifford, R., Rohloff, P., et al. (2019). An Open Source Autocorrelation-Based Method for Fetal Heart Rate Estimation from One-Dimensional Doppler Ultrasound. *Physiol. Meas.* 40, 025005. doi:10.1088/1361-6579/ab033d

Warrick, P. A., and Hamilton, E. F. (2012). Fetal Heart-Rate Variability Response to Uterine Contractions during Labour and Delivery. 2012 Computing in Cardiology, IEEE.

Conflict of Interest: Author PW was employed by the company PeriGen Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pediatrics

BRIEF RESEARCH REPORT
published: 01 September 2021
doi: 10.3389/fped.2021.662101

Fetal Heart Rate Fragmentation

Matilde Costa[1†], *Mariana Xavier*[1†], *Inês Nunes*[2,3,4] *and Teresa S. Henriques*[3,5*]

[1] Department of Biomedical Engineering, Faculty of Engineering, Universidade do Porto, Porto, Portugal, [2] Centro Materno-Infantil do Norte, Centro Hospitalar e Universitário do Porto, Porto, Portugal, [3] Centre for Health Technology and Services Research (CINTESIS), Faculty of Medicine University of Porto, Porto, Portugal, [4] ICBAS School of Medicine and Biomedical Sciences, University of Porto, Porto, Portugal, [5] Department of Health Information and Decision Sciences-MEDCIDS, Faculty of Medicine, University of Porto, Porto, Portugal

Intrapartum fetal monitoring's primary goal is to avoid adverse perinatal outcomes related to hypoxia/acidosis without increasing unnecessary interventions. Recently, a set of indices were proposed as new biomarkers to analyze heart rate (HR), termed HR fragmentation (HRF). In this work, the HRF indices were applied to intrapartum fetal heart rate (FHR) traces to evaluate fetal acidemia. The fragmentation method produces four indices: PIP-Percentage of inflection points; IALS-Inverse of the average length of acceleration/deceleration segments; PSS-Percentage of short segments; PAS-Percentage of alternating segments. On the other hand, the symbolic approach studied the existence of different patterns of length four. We applied the measures to 246 selected FHR recordings sampled at 4 and 2 Hz, where 39 presented umbilical artery's pH $\leq$7.15. When applied to the 4 Hz FHR, the PIP, IASL, and PSS showed significantly higher values in the traces from acidemic fetuses. In comparison, the percentage of "words" W_1^h and W_2^s showed lower values for those traces. Furthermore, when using the 2 Hz, only IASL, W_0, and W_2^m achieved significant differences between traces from both acidemic and normal fetuses. Notwithstanding, the ideal sampling frequency is yet to be established. The fragmentation indices correlated with Sisporto variability measures, especially short-term variability. Accordingly, the fragmentation indices seem to be able to detect pathological patterns in FHR tracings. These indices have the advantage of being suitable and straightforward to apply in real-time analysis. Future studies should combine these indexes with others used successfully to detect fetal hypoxia, improving the power of discrimination in a larger dataset.

Keywords: fetal heart rate, fragmentation, symbolic dynamics, short-term variability, acidemia, umbilical cord pH

OPEN ACCESS

Edited by:
Gianluca Lista, Ospedale dei Bambini Vittore Buzzi, Italy

Reviewed by:
Salvatore Andrea Mastrolia, Ospedale dei Bambini Vittore Buzzi, Italy
Anna Maria Marconi, University of Milan, Italy

***Correspondence:**
Teresa S. Henriques
teresasarhen@med.up.pt

[†] These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *01 April 2021*
Accepted: *13 July 2021*
Published: *01 September 2021*

Citation:
Costa M, Xavier M, Nunes I and Henriques TS (2021) Fetal Heart Rate Fragmentation. Front. Pediatr. 9:662101. doi: 10.3389/fped.2021.662101

1. INTRODUCTION

In the twentieth century, technical advances led to the development of continuous electronic monitoring of fetal heart rate (FHR) and uterine contraction (UC) signals, a technology known as cardiotocography (CTG) (1). This technology constitutes the primary screening method to allow early recognition of fetal distress related to intrapartum fetal hypoxia/ acidosis. Intrapartum fetal monitoring's principal goal is to avoid adverse perinatal outcomes related to hypoxia/acidosis without causing an increase in unnecessary obstetrical interventions, such as cesarean sections or instrumental vaginal deliveries, which are associated with higher maternal and perinatal risks perinatal (2). Intrapartum fetal hypoxia is associated with the lack of an adequate oxygen supply to the fetus, which may lead to metabolic acidosis that, if not reversed, may cause cell dysfunction and

death. The involvement of important fetal organs and systems may cause permanent sequelae, such as hypoxic-ischemic encephalopathy (HIE) in the short-term and cerebral palsy in the long-term or perinatal death. Before labor, on average, the arterial pH of a healthy fetus is around 7.35, whereas, at birth, the average pH of the umbilical artery blood is around 7.25. In this sense, it is considered that moderate neonatal acidosis/acidemia will occur when the pH is, at least, below 7.15 (1).

CTG became widely disseminated in industrialized countries, despite controversial scientific evidence in favor of its routine employment (1). The resulting graph is complex in nature and challenging to interpret. Considerable intra- and interobserver disagreement have been demonstrated in its analysis (3–5), both by inexperienced and experienced healthcare professionals (6–8), which limit CTG sensitivity and specificity. Computer analysis of CTGs was developed to overcome the poor inter and intraobserver agreement on tracing interpretation, to provide an objective evaluation of CTG features that are difficult to assess visually, and also to allow objective quantification of variability (9–11), a parameter that is closely related to the state of fetal oxygenation (12). There are different systems currently available that use different mathematical algorithms to elicit real-time alerts when changes associated with fetal hypoxia are detected (13, 14). Therefore, this is an adjunctive technology to CTG that aims to aid clinicians in the labor ward practice to intervene on time in order to avoid adverse perinatal outcomes related with hypoxia.

Commercially available FHR monitors acquire from Doppler or electrocardiographic signals, beat-to-beat intervals measured in milliseconds, and then convert and round off these values to provide a sequence of instantaneous FHRs, expressed in beats per minute (bpm) (15–17). When data is then exported from the FHR monitors to other devices, it is sampled at 4 Hz (there is an interpolation of signals so that an instantaneous FHR value is provided every 0.25 s) (15–17). Previous studies showed that while the linear time-domain parameters obtained from traces acquired at 2 or 4 Hz are correlated, the similar is not verified when using variability indices and nonlinear parameters, such as entropy (17, 18). In Romagnoli et al. (19) the authors compare several indices from 4 Hz traces and the corresponding down-sampled at 2, 1, 0.4, and 0.2 Hz. A better performance was obtained when using 2 Hz signals.

Recently, Costa et al. (20) proposed a new approach to analyze the heartbeat fragmentation to measure the short-term heart rate variability (STV). The assumption was that pathologic systems manifest the highest degree of heart rate fragmentation. The authors showed that these indices successfully distinguished the heartbeat of normal subjects from those with coronary artery disease. Furthermore, in a subsequent study, Costa et al. (21) introduced a similar approach to the previous analysis but using symbolic dynamics in order to get additional information on the temporal structure of heart rate fragmentation. Modanlou et al., in their study (22), observed that the STV was reduced along with neonatal hypoxemia, while more severe hypoxemia leads to the loss of long–term variability. On the other hand, Druzen et al. (23) showed that fetal hypoxia's early effects increased short and long term variability.

In this work, the new indices of both fragmentation methods were applied to FHR intrapartum traces to detect acidemia, comparing the traces sampling at 4 and 2 Hz.

2. MATERIALS AND METHODS

2.1. Data

The database used in this work is available at Physionet (24)—*CTU-UHB Intrapartum Cardiotocography Database* (25). It contains 552 cardiotocography (CTG) intrapartum recordings with a maximum duration of 90 min each. For this work, it was only selected the last hour of the FHR recordings where the signal loss was lower than 15%. From the 246 selected recordings sampled at 4 Hz, 39 presented the umbilical artery's pH $\leq$ 7.15, which were considered cases of fetal acidemia (pathological). The 2 Hz traces were created, ignoring every other beat of 4 Hz sampling. The main clinical characteristics of the database are summarized in **Table 1**.

2.2. SisPorto

The Omniview-SisPorto system (26, 27) was created for CTG interpretation and analysis, incorporating FIGO 2015 guidelines, in its last version (2). The traces were analyzed using the Omniview SisPorto 4.1 at a sampling frequency of 4 Hz. Four basic CTG features were extracted from the SisPorto analysis:

1. Basal line mean level of the most horizontal and less oscillatory FHR segments, in the absence of fetal movements and uterine contractions, associated with periods of fetal rest;
2. Abnormal short-term variability (STV)—percentage of subsequent FHR signals differing less than 1 bpm;
3. Abnormal long-term variability (LTV)—percentage of FHR signals with a difference between the minimum and maximum values in a 1 min window lower than 5 bpm;
4. Saltatory Index—>35% signals outside a filtered band exceeding 25 bpm in last 30, 20, 10, and 5 min.

2.3. Fragmentation Analysis

Considering the time series $X = \{X_1, X_2, ..., X_N\}$, where X_i represents the time of occurrence of the fetal normal sinus beat in the instance *i*, the differences between consecutive beats were defined as $\Delta X_i = X_i - X_{i-1}$.

2.3.1. Fragmentation Indices

From these time series, four fragmentation indices were computed as proposed by Costa et al. (20). Briefly,

1. PIP-Percentage of inflection points.
   For the calculation of this index, X_i was considered an inflection point when the condition $\Delta X_i * \Delta X_{i+1} \leq 0$ was verified. Furthermore, the points considered could also be divided into two different types of inflection points:

   1.1 PIP*hard*—when $\Delta X_i * \Delta X_{i+1} < 0$.
   1.2 PIP*soft*—when $\Delta X_i * \Delta X_{i+1} = 0$.

   These points represent the instants in which either the acceleration sign inverts (PIP *hard*) or it changes to or from zero (PIP *soft*).

TABLE 1 | Patient and labor characteristics for the CTU-UHB database (25).

med [Q1, Q3]	Normal (n = 207)	Pathologic (n = 39)	Mann-Whitney U-test P-value	Cliff's delta effect size
Gestational age (weeks)	39 [40, 41]	40 [41, 41]	0.007	–0.26 (s)
Weight (grams)	3,075 [3,370, 3,625]	3,225 [3,390, 3,650]	0.336	
Mother age	27 [30, 33]	26 [28, 30]	0.068	0.18 (s)
n (%)			**Fisher Test** P**-value**	
Sex (female)	102 (49%)	19 (49%)	1	
Diabetes	18 (9%)	1 (3%)	0.324	
Hypertension	20 (10%)	2 (5%)	0.543	
Preeclampsia	11 (5%)	1 (3%)	0.697	
Pyrexia	4 (2%)	0	1	
Meconium stained fluid	29 (14%)	4 (10%)	0.797	
Induced labor	95 (46%)	14 (36%)	0.293	
Vaginal delivery	207 (100%)	38 (97%)	0.159	

med, median; Q1, first quartile; Q3, third quartile. s, small effect size.

2. IALS-Inverse of the average length of acceleration and deceleration segments.
   An acceleration or a deceleration can be defined as a segment between two consecutive inflection points in the fetal heart rate. For each segment, if the difference between two beats is negative ($\Delta X_i < 0$) it is considered a deceleration. On the other hand, if the difference is positive it is considered a acceleration ($\Delta X_i > 0$). However, there can also be cases in which $\Delta X_i = 0$, meaning that it is not either an acceleration or a deceleration. For the computation of this parameter, these segments were disregarded.
   The size of each acceleration/deceleration is given by the number of points belonging to X_i within that segment.
3. PSS-Percentage of short segments.
   A short segment is considered short if it contains <3 intervals. The PSS was calculated as the complement of the percentage of points in segments of accelerations or decelerations with three or more intervals. It translates to groups of three or more ΔX_i points with the same negative or positive signals in a row.
4. PAS-Percentage of alternating segments.
   An alternating segment is a sequence of at least four ΔX_i points where the sign differs in every single beat. The PAS measure is looking for the percentage of patterns of accelerations (acc) and deceleration (dec) like "acc-dec-acc-dec" or "dec-acc-dec-acc."

The approach is based on the assumption that the higher the signal's alternation, the more fragmented the time series translates into higher indices.

Figure 1A, shows 50 s (101 points) of a FHR trace sampled at 2 Hz. The trace presents 53 inflection points in which 11 are classified as *hard*. The PIP indices for this trace are the following: $PIP = \frac{53}{101} \approx 53\%$; $PIPhard = \frac{11}{101} \approx 11\%$; $PIPsoft = \frac{42}{101} \approx 42\%$. Also, there are 33 segments between inflection points that are accelerations or decelerations, therefore $IALS = \left[\frac{1+5+1+1+2+1+1+1+2+1+1+4+1+1+3+1+1+3+1+1+1+1+1+1+1+1+1+1+1+2+2+1+3}{33}\right]^{-1} \approx 66\%$. At last, $PSS = 1 - \frac{5+4+3+3+3}{101} = 1 - \frac{18}{101} \approx 82\%$; $PAS = 0$.

2.3.2. Symbolic Fragmentation Indices

The vector $\Delta X_i = X_i - X_{i-1}$ was mapped to a ternary symbolic sequence as follows: $s_i = 0$ if $\Delta X_i = 0$, $s_i = 1$ if $\Delta X_i > 0$, and $s_i = 2$ if $\Delta X_i < 0$. That means that an acceleration corresponded to the number 2, a deceleration corresponded to the number 1, and, in the case of two equal consecutive intervals, it corresponded to 0. Considering i the index of the ternary symbolic sequence, short-terms with 4 elements named "words" (w) were build as follows $w_i = \{s_i, s_{i+1}, ..., s_{i+4-1}\}$.

Transitions from symbol "1" to "2" or vice versa, were termed hard (H) inflection points. Transitions to or from zero were termed soft (S) inflection points. Word groups with only hard, only soft, and a combination of hard and soft inflection points (mixed) were, respectively, labeled W_j^h, W_j^s, and W_j^m, j indicates the number of inflection points. To calculate each word's percentage, we use the total number of each word as denominators.

Figure 1B exhibits 25 s (101 points) of a FHR trace sampled at 4 Hz. Eight words of length four were selected and classified to better illustrate the symbolic fragmentation indices analysis.

2.4. Statistical Analysis

The normality of the fragmentation indices in both groups (normal vs. pathological) was verified by observing the histograms and Q-Q graphs. Since almost all indices' distribution was skewed, values were described with the median and interquartile interval [first quartile-Q1, third quartile-Q3]. The Mann-Whitney U-test was used to compare the indices in each

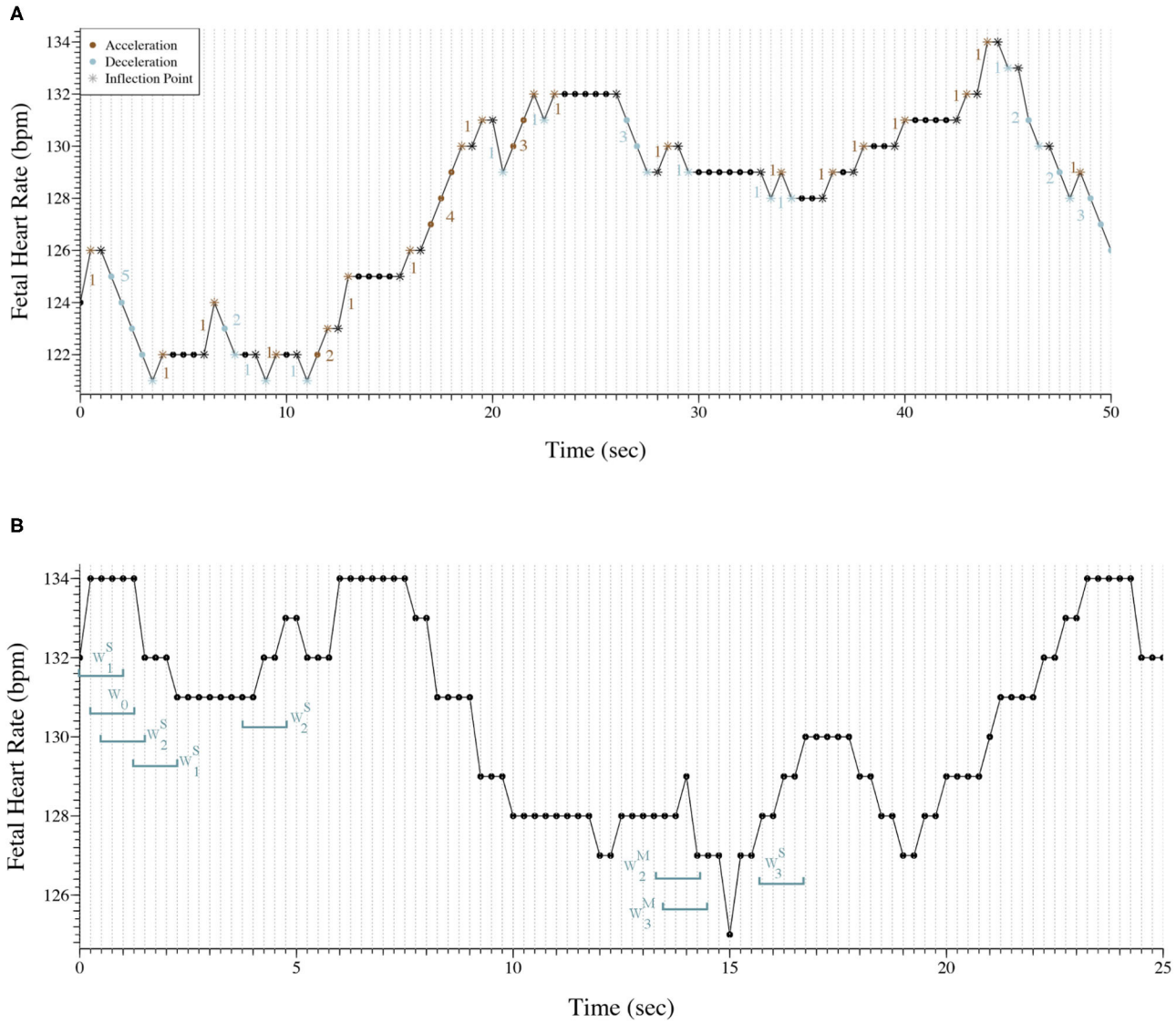

FIGURE 1 | (A) Representation of 50 s of a fetal heart rate time series sampled at 2 Hz. The trace presents 53 inflection points (*) and there are 33 segments between inflections points that are accelerations or decelerations. $PIP = \frac{53}{101} \approx 53\%$; $PIPhard = \frac{11}{101} \approx 11\%$; $PIPsoft = \frac{42}{101} \approx 42\%$; $IALS \approx 66\%$; $PSS \approx 82\%$; $PAS = 0$. **(B)** Representation of 25 s of a fetal heart rate time series sampled at 4 Hz with the classification of eight words as an example.

of the two groups. Cliff's delta was computed to estimate the effect size. Small effect size was considered when Cliff's delta was between 0.15 and 0.33, medium effect size if Cliff's delta was between 0.33 and 0.47 and large effect size when Cliff's delta was higher than 0.47. The correlation between the matching time series' computed indices was calculated using the Spearman correlation coefficients. For descriptive and inference statistics, SPSS Statistics (v.25; IBM SPSS, Chicago, IL) and R software (28) were used. For all statistical tests, it was used a significance level of 0.05.

3. RESULTS

When analyzing the original signals sampled at 4 Hz, from the basic CTG features, only the saltatory index showed significantly higher values in the tracings of the group of fetuses with acidemia compared to those of normal ones. Using the fragmentation measures, we found values of PIP, IASL, and PSS values significantly lower in the tracings of the pathological group (**Table 2**). The higher PIP in the traces from healthy fetuses represent more inflection points, this is, they oscillate more. The lower value of IASL in the tracings of pathological fetuses means that the size of accelerations or decelerations is higher in that group. In agreement with the previously described results, these traces present less beat-to-beat oscillations. Complementary, PSS as a measure of short segments of three or more beats that are not accelerations or decelerations are also lower in the acidemia group. The results also show that the traces analyzed have a large percentage of two consecutive points of equal value (PIP*soft*), which implies a low value of PAS (mostly zeros).

TABLE 2 | Comparison of fragmentation indices for both groups using Mann-Whitney *U*-test when analyzing the 4 Hz fetal heart rate time series.

	Normal		Pathologic			Cliff's delta effect size	Spearman correlation
	med	[Q1–Q3]	med	[Q1–Q3]	*P*-Value		
4 Hz							
Basal line	128	[120–138]	136	[123–142]	0.091	0.17 (s)	
Abnormal STV	46	[39.5–57]	44	[39–55]	0.321		
Abnormal LTV	2	[0–6]	2	[0–5]	0.750		
Saltatory index	79	[48–114]	98	[66.5–134.5]	**0.028**	0.22 (s)	
PIP	97.0	[95.8–97.9]	96.0	[95.0–97.3]	**0.016**	−0.24 (s)	
PIP*hard*	2.5	[1.4–4.2]	3.2	[1.9–4.6]	0.122		
PIP*soft*	94.2	[91.3–96.1]	93.6	[90.9–94.5]	0.056	−0.19 (s)	
IASL	0.92	[0.89–0.94]	0.89	[0.87–0.92]	**0.001**	−0.33 (m)	
PSS	99.4	[98.8–99.8]	99.0	[98.2–99.6]	**0.007**	−0.27 (s)	
PAS	0.00	[0.00–0.03]	0.00	[0.00–0.06]	0.471		
W_0	15.7	[8.9–22.2]	17.0	[10.0–22.2]	0.926		
W_1^s	16.9	[15.1–19.0]	16.7	[14.9–18.4]	0.336		
W_1^h	0.03	[0.01–0.06]	0.06	[0.03–0.11]	**0.003**	0.30 (s)	
W_2^s	31.0	[28.9–32.9]	32.1	[30.8–35.0]	**0.010**	0.26 (s)	
W_2^m	1.9	[1.3–3.3]	2.5	[1.6–3.6]	0.079	0.18 (s)	
W_2^h	0.01	[0.00–0.01]	0.01	[0.00–0.02]	0.277		
W_3^s	21.6	[16.6–30.5]	20.30	[17.0–24.1]	0.290		
W_3^m	8.9	[6.3–11.4]	9.6	[7.5–13.10]	0.141		
W_3^h	0.00	[0.00–0.00]	0.00	[0.00–0.00]	0.136		
2 Hz							
PIP	76.6	[72.3–82.5]	76.8	[72.2–79.9]	0.316		0.33**
PIP*hard*	21.0	[14.8–24.2]	18.6	[14.7– 22.9]	0.262		0.53**
PIP*soft*	58.1	[48.4–65.0]	57.2	[51.1–61.6]	0.657		0.21**
IASL	0.6	[0.6–0.7]	0.6	[0.6–0.7]	**0.046**	−0.26 (s)	0.39**
PSS	79.5	[75.0–85.5]	78.7	[73.1–82.3]	0.091	−0.17 (s)	0.16*
PAS	8.8	[4.5–11.2]	7.8	[4.6–9.4]	0.074	−0.18 (s)	0.29**
W_0	10.3	[7.8–13.0]	11.6	[9.9–13.4]	**0.026**	0.22 (s)	0.71**
W_1^s	16.0	[13.4–21.1]	18.1	[14.3–21.5]	0.260		0.86**
W_1^h	6.2	[4.3–9.3]	6.7	[5.4–10.6]	0.160		0.04
W_2^s	20.1	[18.6–23.2]	21.2	[19.4–22.8]	0.385		0.23**
W_2^m	11.8	[10.7–12.8]	11.0	[10.4–11.8]	**0.024**	−0.23 (s)	0.40**
W_2^h	5.9	[3.6–8.2]	5.1	[4.1–7.1]	0.632		0.32**
W_3^s	7.4	[6.5–8.3]	6.9	[6.0–8.1]	0.120		−0.44**
W_3^m	16.5	[13.0–18.9]	15.1	[12.5–17.2]	0.077	−0.18 (s)	0.58**
W_3^h	3.1	[1.5–4.0]	2.7	[1.4–3.6]	0.194		0.23**

*med, median; Q1, first quartile; Q3, third quartile. s, small effect size; STV, short-term variability; LTV, long-term variability; *p < 0.05, **p < 0.001; m, medium effect size. Bold P-value represent the values lower than 0.05.*

Furthermore, using the symbolic approach in the original traces, the indices W_1^h and W_2^s presented significantly higher values in the tracings of the group of fetuses with acidemia compared to those of normal fetuses. The number of words with soft transitions is much higher than both hard and mixed words. These values corroborate the high percentage of two equal consecutive values.

When the down-sampled 2 Hz signals are analyzed, PAS values increase while PSS values decrease, indicating less repetitive values in the 2 Hz traces than the 4 Hz ones (**Table 2**). However, the hypoxia classification power reduces for all indices. Moreover, the symbolic indices applied to these 2 Hz traces show significantly higher values of W_0 in traces from pathological fetuses, meaning that the traces of pathological fetuses present more patterns of four repeated values than the healthy fetus. Also, we found significantly lower values of W_2^m, meaning that patterns with two inflection points are more frequent in traces from healthy fetuses than pathological ones.

In **Table 2**, the Spearman correlation coefficients between the indices obtained when using the 4 Hz and the matching 2 Hz time series are presented. The achieved correlations are moderate for the fragmentation, being higher for the PIP*hard*

TABLE 3 | Spearman correlation, and corresponding 95% confidence intervals, between Sisporto clinical features, computing using 4 Hz fetal heart rate time series and fragmentation indices for the 4 and 2 Hz fetal heart rate time series.

	Basal line	Abnormal STV	Abnormal LTV	Saltatory index
4 Hz				
PIP	–0.37 [–0.48; –0.25]	0.22 [0.10; 0.34]	0.17 [0.04; 0.29]	–0.33 [–0.44; –0.21]
PIP*hard*	**0.65 [0.57; 0.73]**	0.29 [0.17; 0.40]	0.29 [0.17; 0.40]	–0.11 [–0.24; 0.01]
PIP*soft*	**–0.63 [–0.70; –0.54]**	–0.11 [–0.23; 0.02]	–0.13 [–0.25; 0.00]	–0.05 [–0.18; 0.07]
IASL	–0.34 [–0.45; –0.23]	0.19 [0.07; 0.31]	0.18 [0.06; 0.30]	–0.38 [–0.48; –0.26]
PSS	–0.07 [–0.20; 0.05]	0.22 [0.10; 0.34]	0.17 [0.05; 0.29]	–0.45 [–0.54; –0.33]
PAS	0.29 [0.17; 0.40]	0.06 [–0.07; 0.18]	0.08 [–0.05; 0.20]	0.04 [–0.08; 0.17]
W_0	0.20 [0.08; 0.32]	**0.59 [0.50; 0.67]**	**0.40 [0.29; 0.51]**	–0.36 [–0.47; –0.24]
W_1^s	–0.05 [–0.17; 0.08]	0.36 [0.24; 0.47]	0.09 [–0.04; 0.21]	–0.33 [–0.44; –0.21]
W_1^h	0.02 [–0.11; 0.14]	–0.26 [–0.37; –0.13]	–0.18 [–0.30; –0.05]	**0.45 [0.34; 0.55]**
W_2^s	0.12 [–0.01; 0.24]	**–0.46 [–0.56; –0.36]**	–0.34 [–0.45; –0.22]	**0.48 [0.37; 0.58]**
W_2^m	**0.73 [0.66; 0.79]**	0.43 [0.32; 0.53]	0.38 [0.26; 0.49]	–0.18 [–0.30; –0.05]
W_2^h	0.13 [0.00; 0.25]	–0.12 [–0.25; 0.00]	–0.10 [–0.22; 0.03]	0.27 [0.15; 0.38]
W_3^s	**–0.65 [–0.72; –0.56]**	**–0.62 [–0.70; –0.53]**	–0.44 [–0.54; –0.33]	0.33 [0.21; 0.44]
W_3^m	**0.49 [0.39; 0.59]**	0.05 [–0.08; 0.17]	0.11 [–0.02; 0.23]	0.01 [–0.11; 0.14]
W_3^h	0.29 [0.17; 0.40]	–0.01 [–0.13; 0.12]	0.01 [–0.12; 0.13]	0.13 [0.01; 0.26]
2 Hz				
PIP	0.35 [0.24; 0.46]	**0.86 [0.82; 0.89]**	**0.67 [0.58; 0.74]**	**–0.63 [–0.70; –0.54]**
PIP*hard*	0.02 [–0.11; 0.14]	–0.15 [–0.27; –0.02]	0.00 [–0.12; 0.13]	0.01 [–0.11; 0.14]
PIP*soft*	0.22 [0.09; 0.33]	**0.64 [0.55; 0.71]**	**0.41 [0.30; 0.52]**	**-0.42 [–0.52; –0.31]**
IASL	0.36 [0.24; 0.47]	**0.85 [0.80; 0.88]**	**0.68 [0.60; 0.75]**	**-0.72 [–0.78; –0.65]**
PSS	0.37 [0.25; 0.47]	**0.88 [0.85; 0.91]**	**0.69 [0.60; 0.75]**	**–0.71 [–0.78; –0.64]**
PAS	0.09 [–0.03; 0.22]	0.04 [–0.09; 0.16]	0.12 [0.00; 0.25]	–0.16 [–0.28; –0.03]
W_0	0.20 [0.08; 0.32]	0.22 [0.10; 0.34]	0.15 [0.03; 0.27]	0.06 [–0.07; 0.18]
W_1^s	0.03 [–0.10; 0.15]	0.22 [0.10; 0.34]	0.05 [–0.08; 0.17]	–0.08 [–0.20; 0.05]
W_1^h	–0.30 [–0.42; –0.18]	**–0.80 [–0.85; –0.74]**	**–0.54 [0.50; 0.67]**	**0.60 [0.51; 0.68]**
W_2^s	–0.18[–0.30; –0.05]	0.02 [–0.11; 0.14]	–0.13 [–0.25; –0.01]	–0.01 [–0.14; 0.11]
W_2^m	0.19 [0.07; 0.31]	0.36 [0.24; 0.47]	0.38 [0.26; 0.48]	–0.37 [–0.48; –0.26]
W_2^h	–0.08 [–0.20; 0.05]	–0.36 [–0.47; –0.24]	–0.19 [–0.31; –0.07]	0.18 [0.05; 0.30]
W_3^s	–0.03 [–0.15; 0.10]	0.33 [0.22; 0.44]	0.10 [–0.02; 0.23]	–0.32 [–0.43; –0.20]
W_3^m	0.15 [0.02; 0.27]	0.30 [0.18; 0.41]	0.35 [0.23; 0.46]	–0.32 [–0.43; –0.20]
W_3^h	0.09 [–0.03; 0.22]	0.02 [–0.11; 0.14]	0.13 [0.00; 0.25]	–0.10 [–0.22; 0.02]

STV, Short-term variability; LTV, Long-term variability; bold correlation values represent moderate to high correlations $|r| > 0.40$.

index ($r = 0.53$). Notwithstanding, the values obtained with the 2 Hz time series are almost 10 times higher. In the symbolic fragmentation approach, the words W_0, W_1^s, and W_3^m, exhibited higher correlation values ($r = 0.71$, 0.86, and 0.58, respectively). We highlight the no significant correlation found in the W_1^h index ($r = 0.04$) and the moderate negative correlation obtained in the index W_3^s ($r = -0.44$).

The Spearman correlations between the SisPorto features, computed using the fetal heart rate traces at 4 Hz, and the fragmentation indices at 4 and 2 Hz are presented in **Table 3**. The fragmentation indices PIP*hard* and PIP*soft*, computed in the FHR time series at 4 Hz, are moderately correlated with the basal line values. Moreover, the symbolic fragmentation indices presented moderate correlations with the variability indices. Furthermore, when analyzing the FHR time series at 2 Hz, the PIP, the IASL, the PSS and the W_1^h indices are strongly correlated with the SisPorto variability indices, in particular with the STV. Higher values of the fragmentation indices are correlated with higher values of abnormal STV and LTV. On the other hand, higher values of fragmentation indices represent lower values of Saltatory index.

4. DISCUSSION

The recently proposed fragmentation measures analyze the short-term fluctuations in cardiac beat-to-beat intervals. The novelty of this study is to apply this new fragmentation approach to FHR signals. When applied to FHR, we found that the indices seem to detect pathological patterns in FHR tracings, such as those from acidemic fetuses. In fact, we observed that five of the fragmentation indices, the PIP, IASL, PSS, W_1^h, and the W_2^s, successfully distinguished the traces of fetuses with acidemia from normal fetuses. These indices also have the

advantage of being suitable and straightforward to apply in real-time analysis.

Both fragmentation approaches, the original and the symbolic one, analyze the signal taking into account consecutive accelerations or decelerations, ignoring their magnitude. This procedure relates to the analysis of the STV of the signal. The STV characterizing the beat-to-beat variability is, on average, 2 or 3 bpm and reduced if one or less (1). LTV represents broad-based swings in fetal heart rate, or "waviness," occurring up to several times a minute—it is normal in a bandwidth amplitude of 5–25 bpm. One form of long-term variability of particular significance is a fetal heart "acceleration." These usually occur in response to fetal movement, and are 15 bpm above the baseline or more, lasting 15 s or longer (12). The presence of fetal accelerations is reassuring that the fetus is healthy and tolerating the intra-uterine environment well. Its absence during labor is of no significance. The STV has been studied as one of the predictors of fetal wellbeing in labor, measuring the dynamic interaction between the fetal sympathetic and parasympathetic nervous systems and its effects on fetal cardiovascular activity (29). As the parasympathetic nervous system is more responsible for variations in STV, it might be reduced in central nervous system hypoxia/ acidosis (23). As described before, if hypoxia is sustained and increases in severity, it leads to the loss of long–term variability (22)—resulting in a global decrease of sympathetic and parasympathetic activity. On the other hand, it has been shown that fetal hypoxia's early effects increased short and long term variability (23). The saltatory pattern or increased variability pattern is described as a bandwidth value exceeding 25 bpm lasting more than 30 min (12)—the pathophysiology of this pattern is incompletely understood, but it may be seen linked with recurrent decelerations, when hypoxia/acidosis evolves very rapidly. It is presumed to be caused by fetal autonomic instability/hyperactive autonomic system (30).

Additionally, in the FHR analyzed, the number of consecutive points with the same value is high. The rationale for this finding may be related to the nature of these indices and the redundant values of FHR signals obtained at 4 Hz. A normal fetal heart rate will be expected to vary from 110 to 160 beats-per-minute in an intrapartum setting, corresponding to frequencies between 1.8 and 2.7 Hz. Therefore, we decided to study the indices applied to 2 Hz downsampled time series. Our results verify this theory once correlations were found between the fragmentation indices and the SisPorto variability features. In fact, we found stronger correlations between the fragmentation indices and the SisPorto variability features, especially with the abnormal STV, when considering the 2 Hz time series. In contrast, no correlations were observed if one analyzed the 4 Hz FHR time series.

Our results were consistent with the results observed in previous studies (17–19). We encountered a moderate correlation between the indices computed in the time series with the different sampling rates, but their values varied greatly. In fact, in Romagnoli et al. (19) the authors used the same database used in this paper and considered the 2 Hz acquisition the ideal for their analysis. The results obtained with the 2 Hz seem to be more physiological, but its ability to distinguish traces from acidemic fetuses appears to decrease. A reason for the obtained results might be that when the FHR signal is sampled at 4 Hz when there is no new beat within 0.25 s, a repetition of FHR values will occur, suggesting that 2 Hz sampling may be the best solution. **Figure 1B** is good example of repetitive values in the 4 Hz time series. Almost always, there are at least two consecutive points with the same values. Although, in tachycardia where the FHR increases, more common in the pathologic cases, some information might be lost when using 2 Hz acquisition (17–19).

Furthermore, we believe that the symbolic fragmentation outcomes can be improved. The percentage of non-inflection points might be one of the conditions to be further studied, as well as the length of the word chosen. In the original paper, the choice of words of size four was based on the coupling between the cardio-respiratory systems in adults. In FHR, other sizes should be probed to capture the correct dynamic.

The reduced number of pathologic fetuses limited the number of indices to probe in the logistic regression. Future studies should test the combination of these indexes with others used to detect fetal hypoxia to improve the power of discrimination.

5. CONCLUSION

In this exploratory work, the recently proposed fragmentation measures emerge to detect pathological patterns in FHR tracings. Both fragmentation approaches have the advantage of being quick and straightforward to calculate what may be essential for using these measures in real-time settings. In addition, these measures are related to the Sisporto variability indices, especially with the short-term variability of the signal. The question of the ideal sampling frequency for the FHR time series was raised. If, on one hand, the 2 Hz time series avoid multiple duplicated values, it might lose relevant information when the FHR arises in accelerations and tachycardia episodes. On the other hand, this duality might affect the discriminant power of the indices. Future studies should test the combination of these indexes with others used successfully to detect fetal hypoxia to improve the power of discrimination in a larger dataset. This may contribute to developing new computerized algorithms that may improve CTG diagnostic ability to detect fetal hypoxia/ acidosis.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://physionet.org/content/ctu-uhb-ctgdb/1.0.0/.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements.

AUTHOR CONTRIBUTIONS

MC and MX as Biomedical students analyzed the data, while TH closely supervised the work. TH wrote the paper, conceived and designed the study. IN wrote, edited, and reviewed the clinical part of the manuscript. All the authors contributed to the manuscript draft, revised, read, and approved the final version of the manuscript.

FUNDING

This article was supported by National Funds through FCT–Fundação para a Ciência e a Tecnologia, I.P., within CINTESIS, R&D Unit (reference UIDB/4255/2020).

ACKNOWLEDGMENTS

The authors acknowledge Madalena D. Costa and Ary L. Goldberger for their help. We also acknowledge the SisPorto project based at the department of Obstetrics and Gynecology of the School of Medicine, University of Porto.

REFERENCES

1. Freeman RK, Garite TJ, Nageotte MP, Miller LA. *Fetal Heart Rate Monitoring*. Philadelphia, PA: Lippincott Williams & Wilkins (2012).
2. Ayres-de Campos D, Arulkumaran S, Panel FIFMEC. FIGO consensus guidelines on intrapartum fetal monitoring: physiology of fetal oxygenation and the main goals of intrapartum fetal monitoring. *Int J Gynecol Obstet*. (2015) 131:5–8. doi: 10.1016/j.ijgo.2015.06.018
3. Paneth N, Bommarito M, Stricker J. Electronic fetal monitoring and later outcome. *Obstet Gynecol Survey*. (1994) 49:17–19. doi: 10.1097/00006254-199401000-00009
4. Ayres-de Campos D, Arteiro D, Costa-Santos C, Bernardes J. Knowledge of adverse neonatal outcome alters clinicians' interpretation of the intrapartum cardiotocograph. *BJOG*. (2011) 118:978–84. doi: 10.1111/j.1471-0528.2011.03003.x
5. Ojala K, Mäkikallio K, Haapsamo M, Ijäs H, Tekay A. Interobserver agreement in the assessment of intrapartum automated fetal electrocardiography in singleton pregnancies. *Acta Obstet Gynecol Scand*. (2008) 87:536–40. doi: 10.1080/00016340802023053
6. Doria V, Papageorghiou A, Gustafsson A, Ugwumadu A, Farrer K, Arulkumaran S. Review of the first 1502 cases of ECG-ST waveform analysis during labour in a teaching hospital. *BJOG*. (2007) 114:1202–1207. doi: 10.1111/j.1471-0528.2007.01480.x
7. MacDonald D, Grant A, Sheridan-Pereira M, Boylan P, Chalmers I. The Dublin randomized controlled trial of intrapartum fetal heart rate monitoring. *Am J Obstet Gynecol*. (1985) 152:524–39. doi: 10.1016/0002-9378(85)90619-2
8. Berglund S, Pettersson H, Cnattingius S, Grunewald C. How often is a low Apgar score the result of substandard care during labour? *BJOG*. (2010) 117:968–78. doi: 10.1111/j.1471-0528.2010.02565.x
9. Dawes G, Moulden M, Redman C. Criteria for the design of fetal heart rate analysis systems. *Int J Biomed Comput*. (1990) 25:287–94. doi: 10.1016/0020-7101(90)90032-P
10. Devoe L, Golde S, Kilman Y, Morton D, Shea K, Waller J. A comparison of visual analyses of intrapartum fetal heart rate tracings according to the new national institute of child health and human development guidelines with computer analyses by an automated fetal heart rate monitoring system. *Am J Obstet Gynecol*. (2000) 183:361–6. doi: 10.1067/mob.2000.107665
11. 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) 102:688–700. doi: 10.1111/j.1471-0528.1995.tb11425.x
12. Ayres-de Campos D, Spong CY, Chandraharan E. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynecol Obstet*. (2015) 131:13–24. doi: 10.1016/j.ijgo.2015.06.020
13. Nunes I, Ayres-de Campos D, Figueiredo C, Bernardes J. An overview of central fetal monitoring systems in labour. *J Perinat Med*. (2013) 41:93–9. doi: 10.1515/jpm-2012-0067
14. Nunes I, Ayres-de Campos D. Computer analysis of foetal monitoring signals. *Best Pract Res Clin Obstet Gynaecol*. (2016) 30:68–78. doi: 10.1016/j.bpobgyn.2015.02.009
15. Boos A, Jagger MH, Paret GW, Hausmann JW. A new, lightweight fetal telemetry system. *Hewlett Packard J*. (1995) 46:82–93.
16. Carter MC. Signal processing and display–cardiotocographs. *BJOG*. (1993) 100:21–3. doi: 10.1111/j.1471-0528.1993.tb10631.x
17. Gonçalves H, Costa A, Ayres-de Campos D, Costa-Santos C, Rocha AP, Bernardes J. Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability. *Med Biol Eng Comput*. (2013) 51:665–76. doi: 10.1007/s11517-013-1036-7
18. Gonçalves H, Rocha AP, Ayres-de Campos D, Bernardes J. Internal versus external intrapartum foetal heart rate monitoring: the effect on linear and nonlinear parameters. *Physiol Meas*. (2006) 27:307. doi: 10.1088/0967-3334/27/3/008
19. Romagnoli S, Sbrollini A, Burattini L, Marcantoni I, Morettini M, Burattini L. Digital cardiotocography: what is the optimal sampling frequency? *Biomed Signal Proc Control*. (2019) 51:210–215. doi: 10.1016/j.bspc.2019.02.016
20. Costa MD, Davis RB, Goldberger AL. Heart rate fragmentation: a new approach to the analysis of cardiac interbeat interval dynamics. *Front Physiol*. (2017) 8:255. doi: 10.3389/fphys.2017.00255
21. Costa MD, Davis RB, Goldberger AL. Heart rate fragmentation: a symbolic dynamical approach. *Front Physiol*. (2017) 8:827. doi: 10.3389/fphys.2017.00827
22. Modanlou HD, Freeman RK, Braly P. A simple method of fetal and neonatal heart rate beat-to-beat variability quantitation: preliminary report. *Am J Obstet Gynecol*. (1977) 127:861–868. doi: 10.1016/0002-9378(77)90119-3
23. Druzen M, Ikenoye T, Murata Y, et al. A possible mechanism for the increase in FHR variability following hypoxemia. In: *Presented Before the 26th Annual Meeting of the Society for Gynecologic Investigation.* San Diego (1979).
24. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. (2000) 101:e215–20. doi: 10.1161/01.CIR.101.23.e215
25. Chudáček V, Spilka J, Burša M, Janků P, Hruban L, Huptych M, et al. Open access intrapartum CTG database. *BMC Pregnancy Childbirth*. (2014) 14:1–12. doi: 10.1186/1471-2393-14-16
26. Bernardes J, Moura C, Marques de Sa JP, Pereira Leite L. The Porto system for automated cardiotocographic signal analysis. *J Perinat Med*. (1991) 19:61–5. doi: 10.1515/jpme.1991.19.1-2.61
27. 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) 30:62–7. doi: 10.3109/14767058.2016.1161750
28. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna (2019). Available online at: https://www.R-project.org/.
29. Maulik D, Saini V, Zigrossi S. Clinical significance of short-term variability computed from heart-rate waveforms. *J Perinat Med*. (1983) 11:243–8. doi: 10.1515/jpme.1983.11.5.243
30. Nunes I, Ayres-de Campos D, Kwee A, Rosen K. Prolonged saltatory fetal heart rate pattern leading to newborn metabolic acidosis. *Clin Exp Obstet Gynecol*. (2014) 41:507–11. doi: 10.12891/ceog17322014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Artificial Intelligence

MINI REVIEW
published: 26 October 2021
doi: 10.3389/frai.2021.765210

Challenges of Developing Robust AI for Intrapartum Fetal Heart Rate Monitoring

M. E. O'Sullivan[1], E. C. Considine[1], M. O'Riordan[1,2], W. P. Marnane[1,3], J. M. Rennie[4] and G. B. Boylan[1,5]*

[1]INFANT Research Centre, University College Cork, Cork, Ireland, [2]Department Obstetrics and Gynaecology, University College Cork, Cork, Ireland, [3]School of Engineering, University College Cork, Cork, Ireland, [4]Institute for Women's Health, University College London, London, United Kingdom, [5]Department of Paediatrics and Child Health, University College Cork, Cork, Ireland

Background: CTG remains the only non-invasive tool available to the maternity team for continuous monitoring of fetal well-being during labour. Despite widespread use and investment in staff training, difficulty with CTG interpretation continues to be identified as a problem in cases of fetal hypoxia, which often results in permanent brain injury. Given the recent advances in AI, it is hoped that its application to CTG will offer a better, less subjective and more reliable method of CTG interpretation.

Objectives: This mini-review examines the literature and discusses the impediments to the success of AI application to CTG thus far. Prior randomised control trials (RCTs) of CTG decision support systems are reviewed from technical and clinical perspectives. A selection of novel engineering approaches, not yet validated in RCTs, are also reviewed. The review presents the key challenges that need to be addressed in order to develop a robust AI tool to identify fetal distress in a timely manner so that appropriate intervention can be made.

Results: The decision support systems used in three RCTs were reviewed, summarising the algorithms, the outcomes of the trials and the limitations. Preliminary work suggests that the inclusion of clinical data can improve the performance of AI-assisted CTG. Combined with newer approaches to the classification of traces, this offers promise for rewarding future development.

Keywords: cardiotocography (CTG), fetal heart rate (FHR), hypoxic ischaemic encephalopathy (HIE), labour, pregnancy, fetal hypoxia, artificial intelligence, machine learning

OPEN ACCESS

Edited by:
Antoniya Georgieva, University of Oxford, United Kingdom

Reviewed by:
Philip Warrick, PeriGen Inc., Canada
Danilo Pani, University of Cagliari, Italy

****Correspondence:***
M. E. O'Sullivan
mark.osullivan@ucc.ie

Specialty section:
This article was submitted to Medicine and Public Health, a section of the journal Frontiers in Artificial Intelligence

Received: *26 August 2021*
Accepted: *07 October 2021*
Published: *26 October 2021*

Citation:
O'Sullivan ME, Considine EC, O'Riordan M, Marnane WP, Rennie J and Boylan GB (2021) Challenges of Developing Robust AI for Intrapartum Fetal Heart Rate Monitoring. Front. Artif. Intell. 4:765210. doi: 10.3389/frai.2021.765210

1 INTRODUCTION

Ensuring the safe passage of a baby through the birth canal remains a major challenge globally. Despite improvements in stillbirth and neonatal mortality rates, intrapartum-related hypoxia ("birth asphyxia") is estimated to contribute to almost a quarter of the world's annual 3 million neonatal deaths and almost a half of the 2.6 million third trimester stillbirths (Lee et al., 2013). The WHO estimated in 2005 that as many as 1 million survivors of birth asphyxia may develop cerebral palsy, learning difficulties or other disabilities each year. In England, the 2019/20 annual report of NHS Resolution (NHSR), the body that oversees clinical negligence claims, stated that £2.3 billion was spent on clinical negligence payments, of which 50% went on settling obstetric claims (which

represented just 9% of the total claims made). NHSR estimated that for every baby born in England £1100 was paid in indemnity costs (NHS Resolution, 2019).

Currently, the only non-invasive way of assessing the fetus in labour is by monitoring fetal heartrate. Cardiotocography (CTG) is a technique that measures changes in fetal heart rate (FHR) and relates it to uterine contractions (UC) in order to identify babies who are becoming short of oxygen (hypoxic). CTG monitoring was introduced in the 1960s despite the absence of RCTs. Since then, a Cochrane review of 13 trials involving 37,000 women has shown that continuous CTG monitoring compared to intermittent auscultation was associated with a 50% reduction in neonatal seizures (Alfirevic et al., 2017). The review was dominated by the large Dublin trial which enrolled 12,964 women in 1981–1983 (MacDonald et al., 1985). This trial showed no difference in neonatal mortality or cerebral palsy rates. Many guidelines and textbooks on CTG interpretation have been published over the years, the most recent being the NICE intrapartum care guideline of 2014, updated in 2017 (National Institute for Health and Care Excellence (NICE), 2017). CTG interpretation is heavily dependent on pattern recognition, in particular the FHR response to UCs. Abnormal patterns, such as "late" decelerations, can indicate fetal hypoxia, but the CTG is an overly sensitive test; 60% of babies born after their CTG showed such changes were not acidotic (Beard and Finnegan, 1974). CTG interpretation has low inter- and intra-observer agreement rates, and even experts can differ in their interpretation of the same CTG.

The potential of CTG monitoring has not been realised in spite of major efforts aimed at training staff. NHSR has conducted several reviews (10 years of maternity claims (NHS Resolution, 2018) and 5 years of cerebral palsy claims (NHS Resolution, 2017). Errors with the interpretation of FHR monitoring was the most common theme and were often related to systemic and human factors. Uninterpretable CTGs were also common, with a wait and see approach being taken when there was possible loss of contact. The Royal College of Obstetricians & Gynaecologists (RCOG) "Each Baby Counts" report reached the same conclusion (Royal College of Obstetricians and Gynaecologists, 2020). The latest NHSR review recommended that CTG interpretation should not occur in isolation, but as part of a holistic assessment.

With artificial intelligence (AI), we can now take a fresh, unbiased look at the CTG. Previous attempts at using AI analysis of CTG have not proved successful. Most aimed to mimic human methods of analysis (e.g. recognition of FHR baseline, FHR variability and decelerations). However, modern computer systems using more advanced machine learning methods can include wide ranging analysis. AI systems are available 24/7, and are not affected by human factors such as fatigue, distraction, bias, poor communication, cognitive overload, or fear of doing harm. All of these were identified as limiting factors by the RCOG "Each Baby Counts" reports. Better ways of using and interpreting the CTG have the potential to reduce death and disability, and to prevent significant litigation costs.

2 REVIEW OF PRIOR ART IN AI FOR CTG

2.1 Algorithms Used in Randomised Control Trials

Recent systematic reviews of AI for CTG concluded that prior studies did not manage to improve rates of neonatal acidosis, seizures, death, unnecessary interventions or ICU admissions (Campanile et al., 2018; Balayla and Shrem, 2019; Garcia-Canadilla et al., 2020). One study found that inter-rater reliability between humans and AI was moderate but that AI models that mimic human interpretation is akin to adding a "second evaluator with similar instructions" (Balayla and Shrem, 2019). This suggests that for decision support to be effective, it should add value through features that are not obvious to the human. The three RCTs included in the review paper, which are the only trials that compare human and AI CTG interpretation, are revisited below. The three systems used hand-crafted features that generally aimed to replicate the International Federation of Gynecology and Obstetrics (FIGO) guidelines (Ayres-de-Campos et al., 2015).

The INFANT (Intelligent Fetal AssessmeNT) system was developed over 20 years ago to extract and quantifies the following FHR features: signal quality, baseline, variability, accelerations, decelerations and their timing in relation to contractions. These are the features that are typically interpreted by the human in current clinical practice. The INFANT system extracts these features using numerical algorithms and artificial neural networks (Keith and Greene, 1994). Relevant clinical information, including cervical dilation, analgesia, fetal blood sampling and risk factors (intra-uterine growth restriction, placenta abruption and meconium) are also considered in the AI model. The system uses over 400 rules to interpret the data and provide decision support. It does not provide any recommendations for actions that should be taken in response to detected FHR abnormalities (Keith and Greene, 1994).

A multicentre RCT of this system on 47,000 patients was completed in 2017, which found that the decision-support software did not improve clinical outcomes, despite its effectiveness in correctly detecting FHR abnormalities (Brocklehurst et al., 2017). The hypotheses that substandard care was due to failure to identify non-reassuring CTG and that a decision-support system would reduce unnecessary interventions were not supported. The study suggests that substandard care was due to management decisions after identifying CTG abnormalities. The decision-support system used in the trial did not include clinical information pertaining to the labour (i.e. labour duration and progress). Including this information in the decision support system may have improved decisions to escalate.

Omniview-SisPorto 3.5 provides alerts based on computer analysis of CTG. It classifies CTG into four classes (reassuring, non-reassuring, very non-reassuring and pre-terminal) based on FIGO guidelines (Ayres-de-Campos et al., 2015), including definitions of late/prolonged/repetitive decelerations, reduced variability and baseline variation (Ayres-de-Campos et al.,

2008). Their preliminary results showed that the agreement percentage between human and computer classification of contractions, accelerations and decelerations was 87, 71 and 68%, respectively (Costa et al., 2010). An RCT of the system on 7,320 patients was recently conducted (Nunes et al., 2017). The study concluded that while very low rates of acidosis were observed, the reduction in the rates of acidosis and obstetric interventions between the two arms of the study were not statistically significant.

A smaller RCT was conducted on a quantitative cardiotocography (qCTG) decision-support system, which enrolled 720 patients (Ignatov and Lutomski, 2016). The qCTG system computes features based on three domains: FHR, FHR micro-fluctuations, and decelerations. The features derived from FHR micro-fluctuations are the extrema per minute, the mean beat-to-beat variability per minute, and the oscillation amplitudes. A score of 0–6 is calculated for each domain and summed, giving an overall score of 0–18. The primary outcomes of the trial were hypoxia, acidemia, caesarean section and forceps extraction. Secondary outcomes were 5-minute Apgar, neonatal seizures and NICU admission. Reduced risks were observed for all outcomes in the interventional arm compared to the control arm. However, due to the small sample size of this study, a larger RCT is required to validate these findings.

2.2 Alternative Approaches

The aforementioned RCTs used computer-based algorithms that were largely based on features defined by the thresholds for baseline, variability and decelerations in the FIGO guidelines. Alternative approaches have been investigated which provide AI-based CTG interpretation in a manner that applies feature engineering theory from other domains that may complement existing human interpretation. While such systems have not yet been validated in RCTs, preliminary results are promising.

A control theory approach has been proposed to model the dynamic relationship between FHR and UC as an impulse response function (IRF) (Warrick et al., 2009). Pairing FHR and UC as an input-output system is clinically relevant, as decelerations are classified in response to the contractions. Early decelerations coincide with contractions, and do not indicate fetal hypoxia or acidosis. Late decelerations occur more than 20 s after a contraction and are indicative of hypoxia. Prolonged decelerations span multiple contractions and are indicative of hypoxia (Ayres-de-Campos et al., 2015). This method showed that IRFs in pathological cases resulted in longer delays between contractions and corresponding decelerations. IRF, FHR baseline and FHR variability were used as input features to a support vector machine (SVM) to classify normal and pathological CTGs. The training dataset consisted of 189 normal outcome cases and 31 pathological outcome cases. Their definition of a pathological case was death, or evidence of hypoxic ischemic encephalopathy (HIE), or a base deficit of more than 12 mmol per litre (mmol/L) meaning an acidic pH. The SVM correctly detected 50% of pathological cases with a false positive rate of 7.5% (Warrick et al., 2009).

A method using phase-rectified signal averaging to compute the mean decelerative capacity (DC) of FHR has been proposed (Georgieva et al., 2014). DC was compared to short-term variability (STV), which is considered a strong indicator of pH and has been used in previous studies. The results showed that DC predicted acidemia with 0.665 Area under the Curve (AUC). By comparison, STV achieved 0.606 AUC. Correlation between DC and STV was low, indicating that both may be used in multivariate analysis for improved prediction.

The FHR frequency content can be segmented in to low-frequency (0.04–0.15 Hz), mid-frequency (0.15–0.5 Hz) and high-frequency (0.5–1.0 Hz) bands. These bands correspond to sympathetic activity, fetal movement, and fetal breathing, respectively. The spectral densities and ratios between bands have been previously used to classify normal and pathological CTGs (Signorini et al., 2003; Spilka et al., 2013; Zhao et al., 2018). Fractal analysis and the Hurst parameter have been shown to be a robust alternative to using arbitrarily defined frequency bands, and predicted fetal acidosis with an AUC of 0.87 (Doret et al., 2015).

CTG is a very dynamic signal and the evolution of the CTG toward delivery is significant. An approach described in (Dash et al., 2014) segments the full CTG record into much shorter segments, extracts features and thus represents each full CTG record as a sequence of feature values, which are used as input to a Bayesian classifier. This method achieved a true negative rate (TNR) and true positive rate (TPR) of 0.817 and 0.609, respectively, outperforming SVM models trained on the same dataset.

The aforementioned methods use traditional machine learning, which requires a feature extraction and selection stage before classification. Deep learning is a subset of machine learning, which uses a layered structure of calculations known as neural networks on unstructured data, whereby feature extraction and classification is performed in an optimised end-to-end routine, as depicted in **Figure 1** (Garcia-Canadilla et al., 2020). While deep learning approaches require a relatively larger dataset, it offers the ability to learn complex features from the raw data, which may not be obvious to human experts. Deep Neural Networks (DNNs) were shown to outperform conventional machine learning algorithms, such as SVM and K-Means Clustering, for CTG classification on a database containing 162 normal cases and 162 abnormal cases (defined as pH < 7.20 and/or Apgar at 1 min <7) (Ogasawara et al., 2021). A multi-modal convolutional neural network (MCNN) architecture trained on over 35,000 patients was recently published (Petroziello et al., 2019). The MCNN takes input from the UC, FHR and signal quality measures. Its performance was assessed by measuring the percentage of interventions that were false positives and true positives. A retrospective analysis showed that current clinical practice resulted in a 15% false positive rate (FPR) and a 31% true positive rate (TPR), while the MCNN achieved a 14% FPR and a 50% TPR.

The RCT of the INFANT system concluded that including additional clinical information pertaining to labour could improve outcomes (Brocklehurst et al., 2017). Clinical information including maternal age, prior obstetric outcomes, thick meconium and uterine rupture were shown to be independent risk factors of severe neonatal acidosis

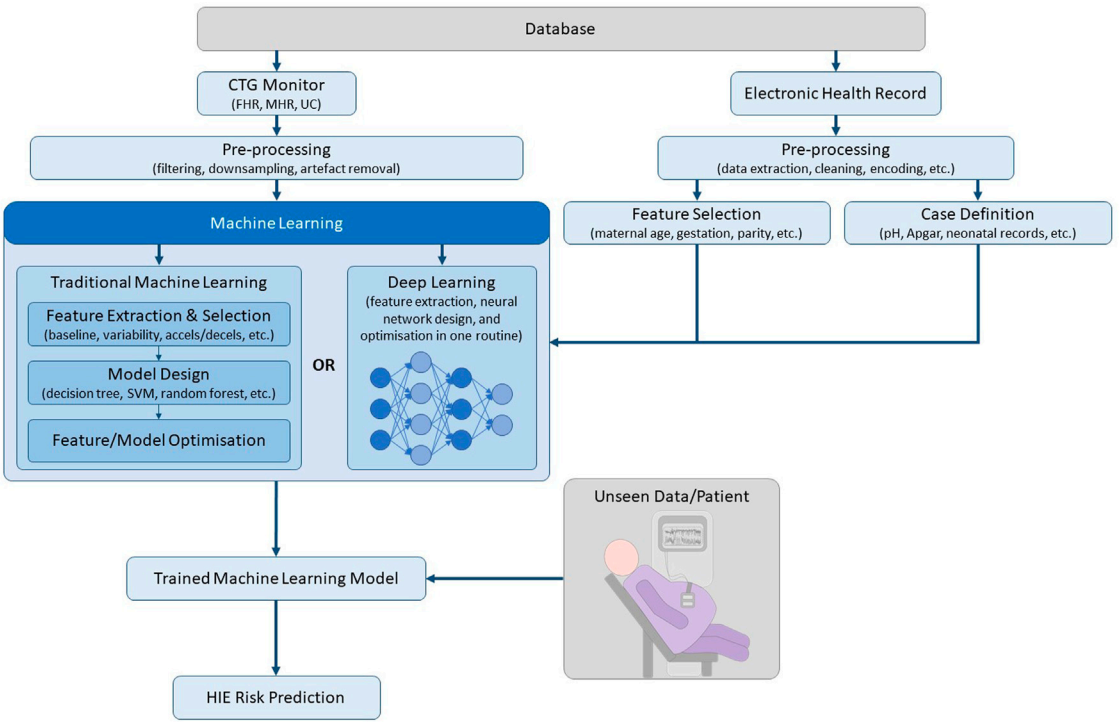

FIGURE 1 | Pipeline of machine learning steps for CTG.

(Maisonneuve et al., 2011). Similarly, results have shown that data-driven systems that use clinical risk factors result in improved classifier performance (Georgieva et al., 2017).

Table 1 summarises the above studies based on method, inputs, target, dataset size, and findings.

3 CHALLENGES IN AI FOR CTG

3.1 Case Definition and Class Imbalance

CTG provides information on how the fetus is coping during labour, with the aim of allowing clinicians to detect non-reassuring fetal status so that adverse outcomes can be avoided through intervention. However, non-reassuring fetal status can result in a spectrum of outcomes, from a wholly unaffected fetus (due to a false positive CTG) to death (Gravett et al., 2016). Therefore, the question arises as to how a "control" patient versus a "pathological case" patient should be labelled in a machine learning architecture.

The incidence rate of HIE is 1-3 per 1,000 in high income countries (Kurinczuk et al., 2010). HIE is the primary condition that a CTG classifier should be trying to predict so that clinicians can intervene and prevent adverse outcome. However, this results in a significant class imbalance between normal and HIE classes, which leads to challenges from a machine learning perspective. At the higher range of 3 per 1,000, it would require over 30,000 deliveries to obtain a database with 100 HIE cases. Minority class oversampling techniques, such as Synthetic Minority Oversampling Technique (SMOTE), have been successfully used in CTG classification studies to introduce synthetic examples in the feature space (Spilka et al., 2013) (Hoodbhoy et al., 2019). However, a sufficient number of genuine cases are still required to use such techniques to synthesize examples. Similarly, weighted errors for misclassifying an example from the minority class has been used to rectify the class imbalance problem (Petroziello et al., 2019).

Due to the difficulty of acquiring a database with comprehensive NICU records and HIE diagnoses, proxy metrics are often used to label classes. There are many proxies for HIE, both objective (pH, base deficit, lactate, and transfer to NICU) and subjective (Apgar scores), with varying degrees of correlation to HIE. Metrics such as pH are generally used as indicators of poor outcome (Malin et al., 2010). However, there is literature that shows ambiguity in the correlation between pH and outcome (Yeh et al., 2012). Quite often, only the umbilical venous pH is measured or recorded, whereas the arterial pH can be significantly lower than the venous pH in babies exposed to a period of acute cord compression shortly before delivery (Westgate et al., 1994). As highlighted in **Table 1**, there is no consistency in the prior art as to what outcome, metric or combination of metrics are used to define a pathological case. A recent systematic review of intrapartum uterine activity and neonatal outcomes found that, of the 12 studies that met the inclusion criteria, 7 used pH as an individual outcome, Apgar scores and base excess were reported as individual outcomes in 4 studies and only 1 study reported neonatal encephalopathy as an outcome. No study examined long-term neurodevelopment as an outcome (Reynolds et al., 2020a). The Apgar score was not

TABLE 1 | Prior art comparison.

Author, year	ML methods	Input features	Target/labels	No. of patients	Type of study	Key finding
Brocklehurst et al. (2017)	Numerical algorithms and artificial neural network	Signal quality, baseline, variability, accels, decels and clinical data (dilation, analgesia, fetal blood sampling, growth restriction, placenta abruption and meconium)	Manually labelled CTG	47,000	RCT	Effective in identifying abnormal CTG, however clinical outcomes not improved
Nunes et al. (2017)		Contractions, accels, decels	Manually labelled CTG	7,320	RCT	Low rates of acidosis observed, however reduction in acidosis between the control arm and the interventional arm were not statistically significant
Ignatov and Lutomski (2016)		FHR, decels, FHR micro-fluctuations (extrema per minute, mean beat-to-beat variability, oscillation amplitudes)	Hypoxia (cord-artery blood ph < 7.20), acidemia (umbilical-artery blood pH < 7.05), intervention (caesarean or forceps)	720	RCT	Reduced risks observed for all targets in interventional arm
Warrick et al. (2009)	Support vector machine	Baseline, variability, impulse response function for decels and contractions	Base deficit (>12 mmol/L), death or HIE	213	Rtrspec. study	50% of pathological cases correctly detected with a false positive rate of 7.5%
Georgieva et al. (2014)		Decelerative capacity	Acidemia (pH < 7.05)	7,568	Rtrspec. study	AUC of 0.665 as a single feature in predicting acidemia
Doret et al. (2015)		Hurst parameter	Acidemia (pH < 7.05)	45	Case control study	AUC of 0.87 in predicting acidosis
Dash et al. (2014)	Generative models and Bayesian theory	FHR baseline, variability, accelerations, decelerations, FHR response to contractions in 4.5–30 mHz, variability in 30–1000 mHz band	Acidemia (ph < 7.15)	83	Rtrspec. study	0.817 TNR and 0.609 TPR
Ogasawara et al. (2021)	CNN	FHR	Acidemia (umbilical artery pH < 7.20) or Apgar at 1 min <7	324	Rtrspec. study	AUC of 0.73 with CNN, which was higher than traditional ML
Petroziello et al. (2019)	Multi-modal CNN	Signal quality, FHR, UC	Acidemia (pH < 7.05) and severe compromise (stillbirth, neonatal death, neonatal encephalopathy, NICU admission)	35,429	Rtrspec. study	Improved prediction of acidemia/compromise compared with clinical practice (14% FPR & 50% TPR versus 15% FPR & 31% TPR)
Georgieva et al. (2017)		Decelerative capacity and clinical data (presence of thick meconium or preeclampsia)	Acidemia (pH < 7.05) and severe compromise (stillbirth, neonatal death, neonatal encephalopathy, NICU admission)	22,790	Cohort study	Improved sensitivity and false-positive rate in detecting acidemia/compromise compared to clinical practice
Hoodbhoy et al. (2019)	XGBoost	21 features including basic quantitative values (max, min, median), STV, and number of fetal movements, decelerations and contractions	Manually labelled CTG	2,126	Rtrspec. study	Overall accuracy of 93%

designed as a measure of birth asphyxia, and a recent cohort study including 85,076 infants concluded that although there is a close association between Apgar score and acidosis, Apgar score should not be used as a measure of birth asphyxia (Cnattingius et al., 2020).

3.2 Weak Labels Versus Expert Annotated Labels

As previously discussed, proxy metrics, such as pH, are often used as individual metrics to distinguish between normal and pathological outcome. CTU-UHB CTG database is a publicly available database hosted on Physionet, which is commonly used for research purposes (cited by over 150 papers) (Chudácek et al., 2014). The database includes 552 CTG recordings from 9,164 recordings acquired from one hospital over a three-year period. Of the 552 patients, 44 had a pH value less than 7.05, which is the threshold commonly used in literature to define pathological cases (Spilka et al., 2013). Annotation by three experts on the same database labelled 149 as normal CTG, 115 as pathological CTG and 275 as suspect CTG. This highlights the disparity between low pH and abnormal CTG (Spilka et al., 2013).

A major challenge with developing machine learning architectures based on proxies and neonatal outcomes is the fact that these labels are "weak." The raw CTG in these cases are not labelled by event or by epoch. Instead, there is one overall label for the patient based on clinical metrics (i.e. pH < 7.05), regardless of the duration of the CTG abnormality, or the type of

hypoxia. Different types of fetal hypoxia (acute, subacute, evolving, chronic) generally manifest in different forms in the CTG, and are associated with widely differing clinical events (Yatham et al., 2020). This introduces problems, as in an acute event (such as cord prolapse, uterine rupture, or acute cord compression) the CTG may only change during the event. Therefore, labelling an entire CTG record as fetal hypoxia may introduce noisy labels and misclassifications. This is particularly problematic if weak labels are being applied to short epochs (i.e. overlapping windows of 15–30 min segments), as there is a significant risk of introducing predominantly noisy labels, unless the fetal hypoxia is chronic and prevalent throughout the duration of the recording. Furthermore, studies have shown that not all infants diagnosed postnatally with HIE have evidence of intrapartum hypoxia in the CTG (using current human interpretation) (Yatham et al., 2020).

Machine learning architectures that use hand-labelled CTG at an event/epoch level by an expert annotator would result in stronger labels and, in theory, achieve improved performance. In light of this, an expert obstetrician has manually labelled the aforementioned CTB-UHB database, which has also been made publicly available to supplement the original database (Romagnoli et al., 2020). Several studies have obtained significantly high percentage agreements between algorithm and human labels (Reynolds et al., 2020b). However, introducing human labels may result in similar clinical outcomes to those observed in the prior RCTs, whereby high algorithm-human agreement is achieved but it is akin to adding a second evaluator with similar instructions. Similarly, multiple studies have shown inter-observer agreement for human CTG interpretation in the range of 30–50% (Yatham et al., 2020) (Hruban et al., 2015) (Rhöse et al., 2014). Therefore, there is a risk that human annotations may introduce human bias into the classification, given that expert use of CTG in general is still widely debated (Garcia-Canadilla et al., 2020).

Classification of CTG at an event level alone, without context of the labour progress and duration is not ideal, as features and patterns that may be considered non-reassuring in 1st stage of labour can be considered normal during the active 2nd stage of labour where contractions become more intense. As the end of the CTG often coincides with the time of birth, it is likely that relevant data pertaining to outcome would be most evident in the later stages of CTG. However, there is considerably more noise and motion artifacts in the later stage. Therefore, classifier performance can vary depending on the stage of labour. Studies have shown that the performance of features for classification of fetal compromise vary significantly as labour progresses (Spilka et al., 2014). As such, many studies in the literature omit 2nd stage data, which may reduce the clinical usefulness of a decision support tool in practice (Spilka et al., 2016).

Having access to large databases, capable of training a deep learning model may help resolve this issue, as the feature extraction and classification process could be completed in an optimized routine. The variation in model performance based on the stage of labour was demonstrated in (Petroziello et al., 2019) using a MCNN trained on 35,000 CTGs. The performance of the MCNN trained on the last 60 min of 1st stage was 0.65 AUC, while the same MCNN model trained on the last 30 min of 2nd stage was 0.71 AUC. The best performance of 0.77 AUC was achieved by training on the last 60 min of CTG, regardless of stage (Petroziello et al., 2019).

4 DISCUSSION

Previous feature-based approaches to automated CTG interpretation that closely follow established CTG clinical guidelines achieve high inter-observer agreement with human interpretation. However, they do not result in improvements in clinical outcomes. The findings of these studies suggest that developing systems to mimic existing guidelines and human interpretation will not improve outcomes. More recent methods, facilitated by more computing power, comprehensive electronic health records, and access to larger datasets have resulted in promising developments. However, these approaches are yet to be validated in a RCT.

The major challenges identified in developing robust AI for CTG interpretation are centred around case definition, labelling and class imbalance, which are inherently linked. The table demonstrates the variability in case definition across the prior art, with many using proxy metrics, such as pH, to label cases as healthy versus HIE. At an incidence rate of 1-3 per 1,000 births, class imbalance is a major concern, and perhaps an anomaly detection approach may be best suited.

While accurately detecting non-reassuring CTG patterns is important, it is not the primary challenge. The primary challenge is determining whether non-reassuring CTG patterns require intervention or not based on the progression of labour and on the risk profile of the mother. Our previous work has demonstrated that improvements in classification performance are achievable by adding both clinical variables (such as gestation, parity and hypertension), as well as duration of labour stages (O'Sullivan et al., 2021). The importance of accurate medical records is critical to the clinical decision-making process. Pre-existing maternal medical conditions such as chronic hypertension, and underlying conditions such as intrauterine growth restriction, can render the utero-placental system more vulnerable to hypoxia during labour (Scheidegger et al., 2019). The clinical team need to consider the risk profile of a pregnancy to aid their assessment of a fetus' tolerance to labour and need to be vigilant for any non-reassuring patterns in high-risk pregnancies. Providing a decision support tool that is developed without consideration of these personalised risk factors and their relationship to neonatal outcomes may result in an increase in unnecessary C-sections and operative delivery rates.

To conclude, there is significant scope and promise for decision support tools in the area of CTG, as demonstrated by prior art. We believe that accurate case definition and

segmentation of the data, combined with the inclusion of pre-existing clinical variables and labour progression data will facilitate the development of an explainable artificial intelligence decision support tool.

AUTHOR CONTRIBUTIONS

MS, GB, JR, and EC contributed to writing the manuscript. WM and MR provided critical review and subject-matter expertise. All authors reviewed the final draft prior to submission.

FUNDING

This work was supported by Science Foundation Ireland (19/FIP/AI/7483).

ACKNOWLEDGMENTS

Thanks to Joye McKernan and Prof. Richard Greene of the National Perinatal Epidemiology Centre, University College Cork for their assistance in data collection throughout the project.

REFERENCES

Alfirevic, Z., Devane, D., Cuthbert, G. M. A., and Devane, D. (2017). Continuous Cardiotocography (CTG) as a Form of Electronic Fetal Monitoring (EFM) for Fetal Assessment during Labour. *Cochrane Database Syst. Rev.* 2, CD006066. doi:10.1002/14651858.CD006066.pub3

Ayres-de-Campos, D., Sousa, P., Costa, A., and Bernardes, J. (2008). Omniview-SisPorto 3.5 - a central Fetal Monitoring Station with Online Alerts Based on Computerized cardiotocogram+ST Event Analysis. *J. Perinat Med.* 36 (3), 260–264. doi:10.1515/JPM.2008.030

Ayres-de-Campos, D., Spong, C. Y., and Chandraharan, E. (2015). FIGO Consensus Guidelines on Intrapartum Fetal Monitoring: Cardiotocography. *Int. J. Gynecol. Obstet.* 131 (1), 13–24. doi:10.1016/j.ijgo.2015.06.020

Balayla, J., and Shrem, G. (2019). Use of Artificial Intelligence (AI) in the Interpretation of Intrapartum Fetal Heart Rate (FHR) Tracings: a Systematic Review and Meta-Analysis. *Arch. Gynecol. Obstet.* 300 (1), 7–14. doi:10.1007/s00404-019-05151-7

Beard, R. W., and Finnegan, T. S. (1974). *Fetal Heart Rate Patterns and Their Clinical Interpretation*. Sussex: Sonicaid Ltd.

Brocklehurst, P., Field, D., Greene, K., Juszczak, E., Keith, R., Kenyon, S., et al. (2017). Computerised Interpretation of Fetal Heart Rate during Labour (INFANT): a Randomised Controlled Trial. *The Lancet* 389 (10080), 1719–1729. doi:10.1016/s0140-6736(17)30568-8

Campanile, M., D'Alessandro, P., Della Corte, L., Saccone, G., Tagliaferri, S., Arduino, B., et al. (2018). Intrapartum Cardiotocography with and without Computer Analysis: a Systematic Review and Meta-Analysis of Randomized Controlled Trials. *J. Maternal-Fetal Neonatal Med.* 33 (13), 2284–2290. doi:10.1080/14767058.2018.1542676

Chudácek, V., Spilka, J., Bursa, M., Janku, P., Hruban, L., Huptych, M., et al. (2014). Open Access Intrapartum CTG Database. *BMC Pregnancy Childbirth* 14 (16), 1–12. doi:10.1186/1471-2393-14-16

Cnattingius, S., Johansson, S., and Razaz, N. (2020). Rates of Metabolic Acidosis at Birth and Apgar Score Values at 1, 5, and 10 Min in Term Infants: a Swedish Cohort Study. *J. Perinatal Med.* 48 (5), 514–515. doi:10.1515/jpm-2019-0429

Costa, M. A., Ayres-de-Campos, D., Machado, A. P., Santos, C. C., and Bernardes, J. (2010). Comparison of a Computer System Evaluation of Intrapartum Cardiotocographic Events and a Consensus of Clinicians. *J. Perinat Med.* 38 (2), 191–195. doi:10.1515/jpm.2010.030

Dash, S., Quirk, J. G., and Djuric, P. M. (2014). Fetal Heart Rate Classification Using Generative Models. *IEEE Trans. Biomed. Eng.* 61 (11), 2796–2805. doi:10.1109/tbme.2014.2330556

Doret, M., Spilka, J., Chudáček, V., Gonçalves, P. P., and Abry, P. (2015). Fractal Analysis and Hurst Parameter for Intrapartum Fetal Heart Rate Variability Analysis: A Versatile Alternative to Frequency Bands and LF/HF Ratio. *PLoS One* 10 (8), e0136661. doi:10.1371/journal.pone.0136661

Garcia-Canadilla, P., Sanchez-Martinez, S., Crispi, F., and Bijnens, B. (2020). Machine Learning in Fetal Cardiology: What to Expect. *Fetal Diagn. Ther.* 47, 363–372. doi:10.1159/000505021

Georgieva, A., Papageorghiou, A., Payne, S., Moulden, M., and Redman, C. (2014). Phase-rectified Signal Averaging for Intrapartum Electronic Fetal Heart Rate Monitoring Is Related to Acidaemia at Birth. *Bjog: Int. J. Obstet. Gy* 121 (7), 889–894. doi:10.1111/1471-0528.12568

Georgieva, A., Redman, C. W. G., and Papageorghiou, A. T. (2017). Computerized Data-Driven Interpretation of the Intrapartum Cardiotocogram: a Cohort Study. *Acta Obstet. Gynecol. Scand.* 96 (7), 883–891. doi:10.1111/aogs.13136

Gravett, C., Eckert, L. O., Gravett, M. G., Dudley, D. J., Stringer, E. M., Mujobu, T. B. M., et al. (2016). Non-reassuring Fetal Status: Case Definition & Guidelines for Data Collection, Analysis, and Presentation of Immunization Safety Data. *Vaccine* 34 (49), 6084–6092. doi:10.1016/j.vaccine.2016.03.043

Hoodbhoy, Z., Noman, M., Shafique, A., Nasim, A., Chowdhury, D., and Hasan, B. (2019). Use of Machine Learning Algorithms for Prediction of Fetal Risk Using Cardiotocographic Data. *Int. J. Appl. Basic Med. Res.* 9 (4), 226–230. doi:10.4103/ijabmr.IJABMR_370_18

Hruban, L., Spilka, J., Chudáček, V., Janků, P., Huptych, M., Burša, M., et al. (2015). Agreement on Intrapartum Cardiotocogram Recordings between Expert Obstetricians. *J. Eval. Clin. Pract.* 21 (4), 694–702. doi:10.1111/jep.12368

Ignatov, P. N., and Lutomski, J. E. (2016). Quantitative Cardiotocography to Improve Fetal Assessment during Labor: a Preliminary Randomized Controlled Trial. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 205, 91–97. doi:10.1016/j.ejogrb.2016.08.023

Keith, R. D. F., and Greene, K. R. (1994). 4 Development, Evaluation and Validation of an Intelligent System for the Management of Labour. *Baillière's Clin. Obstet. Gynaecol.* 8 (3), 583–605. doi:10.1016/s0950-3552(05)80200-7

Kurinczuk, J. J., White-Koning, M., and Badawi, N. (2010). Epidemiology of Neonatal Encephalopathy and Hypoxic-Ischaemic Encephalopathy. *Early Hum. Develop.* 86 (6), 329–338. doi:10.1016/j.earlhumdev.2010.05.010

Lee, A. C., Kozuki, N., Blencowe, H., Vos, T., Bahalim, A., Darmstadt, G. L., et al. (2013). Intrapartum-related Neonatal Encephalopathy Incidence and Impairment at Regional and Global Levels for 2010 with Trends from 1990. *Pediatr. Res.* 74, 50–72. doi:10.1038/pr.2013.206

MacDonald, D., Grant, A., Sheridan-Pereira, M., Boylan, P., and Chalmers, I. (1985). The Dublin Randomized Controlled Trial of Intrapartum Fetal Heart Rate Monitoring. *Am. J. Obstet. Gynecol.* 152 (5), 524–539. doi:10.1016/0002-9378(85)90619-2

Maisonneuve, E., Audibert, F., Guilbaud, L., Lathelize, J., Jousse, M., Pierre, F., et al. (2011). Risk Factors for Severe Neonatal Acidosis. *Obstet. Gynecol.* 118 (4), 818–823. doi:10.1097/aog.0b013e31822c9198

Malin, G. L., Morris, R. K., and Khan, K. S. (2010). Strength of Association between Umbilical Cord pH and Perinatal and Long Term Outcomes: Systematic Review and Meta-Analysis. *BMJ* 340, c1471–13. doi:10.1136/bmj.c1471

National Institute for Health and Care Excellence (NICE) (2017). Intrapartum Care for Healthy Women and Babies (CG190). [Online]. Available: https://www.nice.org.uk/guidance/cg190.

NHS Resolution (2017). A Summary of: Five Years of Cerebral Palsy Claims: A Thematic Review of NHS Resolution Data. [Online]. Available: https://resolution.nhs.uk/resources/five-years-of-cerebral-palsy-claims/ (Accessed May 13, 2021).

NHS Resolution (2019). Annual Report and Accounts 2019/20. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/901085/nhs-resolution-2019-20-annual-report-and-accounts.pdf (Accessed May 13, 2021).

NHS Resolution (2018). Ten Years of Maternity Claims: An Analysis of NHS Litigation Authority Data. [Online]. Available: https://resolution.nhs.uk/

resources/ten-years-of-maternity-claims-an-analysis-of-nhs-litigation-authority-data/ (Accessed May 13, 2021).

Nunes, I., Ayres-de-Campos, D., Ugwumadu, A., Amin, P., Banfield, P., Nicoll, A., et al. (2017). Central Fetal Monitoring with and without Computer Analysis. *Obstet. Gynecol.* 129 (1), 83–90. doi:10.1097/aog.0000000000001799

O'Sullivan, M., Gabruseva, T., Boylan, G. B., O'Riordan, M., Lightbody, G., and Marnane, W. (2021). "Classification of Fetal Compromise during Labour: Signal Processing and Feature Engineering of the Cardiotocograph"in European Signal Processing Conference, Dublin, Ireland, August 23–27, 2021. Dublin: EUSIPCO).

Ogasawara, J., Ikenoue, S., Yamamoto, H., Sato, M., Kasuga, Y., Mitsukura, Y., et al. (2021). Deep Neural Network-Based Classification of Cardiotocograms Outperformed Conventional Algorithms. *Scientific Rep.* 11–13367. doi:10.1038/s41598-021-92805-9

Petroziello, A., Redman, C. W., Papageorghiou, A. T., Jordanov, I., and Georgieva, A. (2019). Multimodal Convolutional Neural Networks to Detect Fetal Compromise during Labor and Delivery. *IEEE Access* 7, 112026–112036. doi:10.1109/access.2019.2933368

Reynolds, A. J., Geary, M. P., and Hayes, B. C. (2020). Intrapartum Uterine Activity and Neonatal Outcomes: a Systematic Review. *BMC Pregnancy Childbirth* 20 (532), 532. doi:10.1186/s12884-020-03219-w

Reynolds, A. J., Waldron, O. M., Halpern, E. M., McGarvey, C. M., Murray, M. L., Ater, S. B., et al. (2020). A Wavelet-Based Algorithm for Automated Analysis of External Tocography: How Does it Compare to Human Interpretation. *Comput. Biol. Med.* 122, 103814. doi:10.1016/j.compbiomed.2020.103814

Rhöse, S., Heinis, A. M. F., Vandenbussche, F., van Drongelen, J., and van Dillen, J. (2014). Inter- and Intra-observer Agreement of Non-reassuring Cardiotocography Analysis and Subsequent Clinical Management. *Acta Obstet. Gynecol. Scand.* 93 (6), 596–602. doi:10.1111/aogs.12371

Romagnoli, S., Sbrollini, A., Burattini, L., Marcantoni, I., Morettini, M., and Burattini, L. (2020). Annotation Dataset of the Cardiotocographic Recordings Constituting the "CTU-CHB Intra-partum CTG Database". *Data Brief* 31, 105690. doi:10.1016/j.dib.2020.105690

Royal College of Obstetricians and Gynaecologists (2020) Each Baby Counts: 2020 Final Progress Report. [Online]. Available: https://www.rcog.org.uk/en/guidelines-research-services/audit-quality-improvement/each-baby-counts/reports-updates/2020-report/ (Accessed SMarch 18, 2021).

Scheidegger, S., Held, U., Grass, B., Latal, B., Hagmann, C., and Brotschi, B. (2019). Association of Perinatal Risk Factors with Neurological Outcome in Neonates with Hypoxic Ischemic Encephalopathy. *J. Maternal-Fetal Neonatal Med.* 34 (7), 1020–1027. doi:10.1080/14767058.2019.1623196

Signorini, M. G., Magenes, G., Cerutti, S., and Arduini, D. (2003). Linear and Nonlinear Parameters for the Analysis of Fetal Heart Rate Signal from Cardiotocographic Recordings. *IEEE Trans. Biomed. Eng.* 50 (3), 365–374. doi:10.1109/tbme.2003.808824

Spilka, J., Frecon, J., Leonarduzzi, R., Pustelnik, N., Abry, P., and Doret, M. (2016). Sparse Support Vector Machine for Intrapartum Fetal Heart Rate Classification. Computing in Cardiology 2014, Cambridge, MA, September 7–10, 2014 21 (3), 664–671. doi:10.1109/JBHI.2016.2546312

Spilka, J., Abry, P., Goncalves, P., and Doret, M. (2014). "Impacts of First and Second Labour Stages on Hurst Parameter Based Intrapartum Fetal Heart Rate Analysis," in Computing in Cardiology 2014, Cambridge, MA, September 7–10, 2014 (IEEE).

Spilka, J., Georgoulas, G., Karvelis, P., Oikonomou, V. P., Chudáček, V., Stylios, C., et al. (2013). Automatic Evaluation of FHR Recordings from CTU-UHB CTG Database. *Inf. Tech. Bio- Med. Inform.* 8060, 47–61. doi:10.1007/978-3-642-40093-3_4

Warrick, P. A., Hamilton, E. F., Precup, D., and Kearney, R. E. (2009). Identification of the Dynamic Relationship between Intrapartum Uterine Pressure and Fetal Heart Rate for normal and Hypoxic Fetuses. *IEEE Trans. Biomed. Eng.* 56 (6), 1587–1597. doi:10.1109/tbme.2009.2014878

Westgate, J., Garibaldi, J. M., and Greene, K. R. (1994). Umbilical Cord Blood Gas Analysis at Delivery: a Time for Quality Data. *BJOG:An Int. J. O&G* 101 (12), 1054–1063. doi:10.1111/j.1471-0528.1994.tb13581.x

Yatham, S. S., Whelehan, V., Archer, A., and Chandraharan, E. (2020). Types of Intrapartum Hypoxia on the Cardiotocograph (CTG): Do They Have Any Relationship with the Type of Brain Injury in the MRI Scan in Term Babies? *J. Obstet. Gynaecol.* 40 (5), 688–693. doi:10.1080/01443615.2019.1652576

Yeh, P., Emary, K., and Impey, L. (2012). The Relationship between Umbilical Cord Arterial pH and Serious Adverse Neonatal Outcome: Analysis of 51 519 Consecutive Validated Samples. *BJOG Int. J. Obstet. Gynaecol.* 119 (7), 824–831. doi:10.1111/j.1471-0528.2012.03335.x

Zhao, Z., Zhang, Y., and Deng, Y. (2018). A Comprehensive Feature Analysis of the Fetal Heart Rate Signal for the Intelligent Assessment of Fetal State. *Jcm* 7 (8), 223–243. doi:10.3390/jcm7080223

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Medicine

ORIGINAL RESEARCH
published: 25 November 2021
doi: 10.3389/fmed.2021.626450

Relationship Between Deceleration Morphology and Phase Rectified Signal Averaging-Based Parameters During Labor

Massimo W. Rivolta[1], Moira Barbieri[2], Tamara Stampalija[2,3], Roberto Sassi[1] and Martin G. Frasch[4]*

[1] Dipartimento di Informatica, Università degli Studi di Milano, Milan, Italy, [2] Unit of Fetal Medicine and Prenatal Diagnosis, Institute for Maternal and Child Health IRCCS Burlo Garofolo, Trieste, Italy, [3] Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste, Italy, [4] Department of Obstetrics and Gynecology and Center on Human Development and Disability (CHDD), School of Medicine, University of Washington, Seattle, WA, United States

During labor, uterine contractions trigger the response of the autonomic nervous system (ANS) of the fetus, producing sawtooth-like decelerations in the fetal heart rate (FHR) series. Under chronic hypoxia, ANS is known to regulate FHR differently with respect to healthy fetuses. In this study, we hypothesized that such different ANS regulation might also lead to a change in the FHR deceleration morphology. The hypothesis was tested in an animal model comprising nine normoxic and five chronically hypoxic fetuses that underwent a protocol of umbilical cord occlusions (UCOs). Deceleration morphologies in the fetal inter-beat time interval (FRR) series were modeled using a trapezoid with four parameters, i.e., baseline b, deceleration depth a, UCO response time τ_u and recovery time τ_r. Comparing normoxic and hypoxic sheep, we found a clear difference for τ_u (24.8 ± 9.4 vs. 39.8 ± 9.7 s; $p < 0.05$), a (268.1 ± 109.5 vs. 373.0 ± 46.0 ms; $p < 0.1$) and $\Delta\tau = \tau_u - \tau_r$ (13.2 ± 6.9 vs. 23.9 ± 7.5 s; $p < 0.05$). Therefore, the animal model supported the hypothesis that hypoxic fetuses have a longer response time τ_u and larger asymmetry $\Delta\tau$ as a response to UCOs. Assessing these morphological parameters during labor is challenging due to non-stationarity, phase desynchronization and noise. For this reason, in the second part of the study, we quantified whether acceleration capacity (AC), deceleration capacity (DC), and deceleration reserve (DR), computed through Phase-Rectified Signal Averaging (PRSA, known to be robust to noise), were correlated with the morphological parameters. DC, AC and DR were correlated with τ_u, τ_r and $\Delta\tau$ for a wide range of the PRSA parameter T (Pearson's correlation $\rho > 0.8$, $p < 0.05$). In conclusion, deceleration morphologies have been found to differ between normoxic and hypoxic sheep fetuses during UCOs. The same difference can be assessed through PRSA based parameters, further motivating future investigations on the translational potential of this methodology on human data.

Keywords: phase-rectified signal averaging (PRSA), animal model, fetal heart rate (FHR), electronic fetal monitoring (EFM), heart rate variability (HRV), fetal hypoxia, labor

OPEN ACCESS

Edited by:
Mark I. Evans, Icahn School of Medicine at Mount Sinai, United States

Reviewed by:
Serena Xodo, University of Udine, Italy
Tanja Premru-Sršen, University Medical Centre Ljubljana, Slovenia

****Correspondence:***
Massimo W. Rivolta
massimo.rivolta@unimi.it

Specialty section:
This article was submitted to Obstetrics and Gynecology, a section of the journal Frontiers in Medicine

Received: *05 November 2020*
Accepted: *31 October 2021*
Published: *25 November 2021*

Citation:
Rivolta MW, Barbieri M, Stampalija T, Sassi R and Frasch MG (2021) Relationship Between Deceleration Morphology and Phase Rectified Signal Averaging-Based Parameters During Labor. Front. Med. 8:626450. doi: 10.3389/fmed.2021.626450

1. INTRODUCTION

During labor, a fetus might suffer considerable stress due to uterine contractions, causing transient oxygen reduction and head compression, resulting in vagal and sympathetic stimulations. Nutrient deprivation, hypoxemia, hypoxia, acidemia and cardiovascular decompensation directly impact the autonomic nervous system (ANS) and thus affect the fetal heart rate variability (FHRV) (1, 2). The cardiotocography (CTG) remains the best available proxy of ANS' functional state through the analysis of fetal heart rate (FHR) and its variability. Considering that the standard processing of CTG series has been found poorly correlated to the relevant clinical outcomes, such as fetal brain injury or death, new FHR biomarkers are needed to better quantify the risk of fetal morbidity and mortality during labor (3, 4).

Phase-Rectified Signal Averaging (PRSA) analysis extracts quasi-periodic oscillations from HRV series and it is more resistant to non-stationarities, signal loss and artifacts (5) than conventional HRV analysis techniques, such as the well-known spectral analysis. It provides two measures that quantify the average cardiac acceleration (AC) and deceleration (DC) capacities from an inter-beat time interval series (RR). Practically, these measures quantify the average RR increase (or decrease) in milliseconds. When quantified on CTG signals or fetal RR series (FRR), AC and DC seem to perform better than the short term variation of FHR in identifying fetal growth restricted fetuses (6–8) and adverse outcome (9). In a study of fetal sheep exposed to repetitive umbilical cord occlusions (UCOs), a model of uterine contractions during labor, we found that there was a high correlation between AC and DC and acid/base balance (10); particularly, AC and DC progressively increased with the severity of the UCOs, suggesting an activation of ANS of healthy normoxic fetus exposed to acute hypoxemia.

In the same animal model, we recently observed that, at the beginning of each UCO, FRR adapted by a progressive increase (reduction in FHR) and quickly recovered when pressure was released. In order to quantify such adaptations, we modeled the FRR deceleration using a first-order exponential model, one of the possible models typically employed for system modeling tasks (11), for both response and recovery phases (10). These models were characterized by time constants, describing the speed of FRR adaptation (the larger the time constant, the slower the adaptation) and we found that healthy normoxic fetuses had longer UCO response times than the time necessary to return to the baseline level (10), suggesting the presence of asymmetric trends in the series during labor.

Motivated by this observation, we also proved that dissimilarities in AC and DC values arise when asymmetric increasing/decreasing trends appear in the series (12), which seem to occur during labor. We thus introduced the deceleration reserve (DR), a new PRSA-based metric for the quantification of such asymmetry (12). The DR is computed as the difference between DC and AC. Up to date, DR was tested on a near-term pregnant sheep model and human CTG recordings, obtaining promising results for distinguishing between normoxic and chronically hypoxic fetuses, and to detect fetal acidemia at birth. Even though PRSA processes the FRR series in its whole entirety, it is reasonable to hypothesize that AC, DC and DR are deeply linked with the FRR adaptation time due to uterine contractions during labor.

In this study, we hypothesized that the different ANS regulation under chronic hypoxia might also lead to a change in the FHR deceleration morphology, as a result of the uterine contraction, and that the adaptation times would be different from those of healthy fetuses. The hypothesis was tested in an animal model comprising nine normoxic and five chronically hypoxic fetuses that underwent a protocol of UCOs. Deceleration morphologies on the FRR series were modeled using a trapezoid with four parameters characterizing the adaptation times, baseline and deceleration depth. The parameters were compared between the two groups. We also quantified their correlation with biomarkers of acid/base balance. Assessing these morphological parameters during labor is challenging due to non-stationarity, phase desynchronization and noise. For this reason, in the second part of the study, we quantified whether AC, DC and DR were correlated with the morphological parameters. Given the fact that PRSA is more robust with respect to phase-desynchronization, a correlation might further support the opportunity of using AC, DC and DR in the clinical settings. In addition, given the link between the deceleration morphology and the area under the deceleration, *i.e.*, the so-called "Deceleration Area" (DA), we quantified the relationship between PRSA-based parameters and DA.

2. MATERIALS AND METHODS

2.1. Animal Model and FHR Data

An established pregnant sheep model of labor was retrospectively analyzed. A comprehensive review on the pregnant sheep model and its translational significance for human physiology, in particular for studies of the ANS, can be found in Morrison et al. (13). The animal cohort comprised of nine normoxic and five spontaneously chronically hypoxic near-term pregnant sheep fetuses which underwent periodic UCOs mimicking uterine contractions during labor.

The animal and experimental models were described elsewhere (14). The animal study was reviewed and approved by University of Western Ontario Council on Animal Care/Canadian Council on Animal Care. Animal care followed the guidelines of the Canadian Council on Animal Care and was approved by the University of Western Ontario Council on Animal Care. Briefly, sheep fetuses were monitored over a 6 h period during which a mechanical stimulation was applied to the umbilical cord by using an inflatable silicon rubber cuff. A baseline period of approximately 1 h with no occlusion preceded the study. After that, UCOs were delivered every 2.5 min and lasted for 1 min. Three levels of occlusion strength, from partial to complete, were performed: mild (MILD, 60 min), moderate (MODERATE, 60 min) and complete (SEVERE, 60 min or until pH $<$ 7.00 was reached). The stimulation protocol ended with a recovery period. During the stimulation protocol, pH, base deficit (BE) and lactate (hereafter referred to as "biomarkers")

were quantified by means of fetal arterial blood samples collected every 20 min.

Sheep fetuses were categorized as chronically hypoxic if $O_2Sat < 55\%$, as measured before the beginning of the UCO stimulation protocol. In this study, we refer to the two models as "normoxic" and "chronically hypoxic," respectively. As per experimental protocol, both models showed a progressive worsening acidemia of the hypoxic status until pH < 7.00 was reached [see **Figure 2** in (12)].

Fetal ECGs were collected by means of electrodes implanted into the left supra-scapular muscles, in the muscles of the right shoulder and in the cartilage of the sternum, and digitized at 1,000 Hz. FRR series were automatically extracted from the fetal ECG (15).

In this study, we only considered the SEVERE phase of UCOs since FHR mostly changed during this condition.

2.2. FHR Series Preprocessing

A preprocessing similar to the one proposed in Rivolta et al. (12) was adopted for both datasets. Briefly, FRR series were analyzed to determine whether they were suitable for further analysis in terms of noise level, by excluding those recordings with more than 10% gaps during the SEVERE phases. Two normoxic fetuses were excluded from the analysis because of the high amount of missing beats. Furthermore, FRR intervals greater than 1,500 ms (40 bpm) were labeled as artifacts and substituted with an equivalent number of beats (calculated dividing the length of each artifact by the median of the 20 nearby FRR samples). The reconstructed samples were used neither in the model fitting nor as anchor points in the PRSA analysis (in this latter case, however, they contributed to the selection of nearby anchor points).

2.3. Geometrical Model of Deceleration Morphology and Its Fitting to FHR Data

In this analysis, we quantified the average characteristics of the FHR response to UCOs in terms of baseline level, deceleration depth, time necessary to reach a steady condition of the FRR during both UCO stimulation and resting phase in the normoxic and hypoxic datasets. To do so, we used a simple model to describe the time evolution of FRR during each cycle of UCO and rest and extract the relevant information. The procedure took two steps. First, we time-aligned all the FRR segments of 150 s starting from the beginning of each UCO. Second, a piecewise linear model was fitted using a semi-automatic approach based on least squares. The model was as follows during UCO stimulation

$$y_u(t) = b + \begin{cases} 0 & t < 0 \\ \frac{a}{\tau_u} t & 0 \le t < \tau_u \\ a & \tau_u \le t < 60 \end{cases} \quad (1)$$

and the following one for the resting phase

$$y_r(t) = b + \begin{cases} a - \frac{a}{\tau_r}(t - 60) & 60 \le t < 60 + \tau_r \\ 0 & 60 + \tau_r \le t < 150 \end{cases} \quad (2)$$

where t was the time (seconds), b (milliseconds), a (milliseconds), τ_r (seconds), and τ_u (seconds) were the morphological parameters to be estimated. In particular, b was the baseline FRR value in absence of UCOs, τ_u the time to reach the steady condition during UCO, a is the amplitude change of FRR, and τ_r the time to reach the baseline b after releasing the UCO. In addition, the difference $\Delta\tau = \tau_u - \tau_r$ was considered as measure of asymmetry to the response to UCO stimulation. **Figure 1** shows an example of model fitting and a visual description of the morphological parameters.

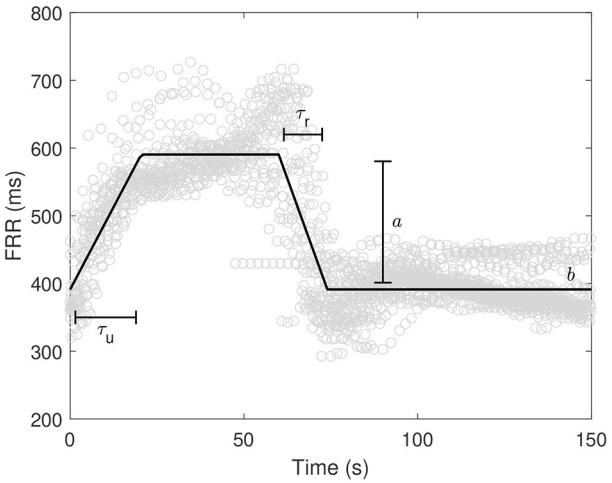

FIGURE 1 | Example of model fitting along with variable definition. ◦ refers to FRR values from the beginning of the UCOs.

2.4. Correlation of Deceleration Morphology With Time and Biomarkers

In the second analysis, we determined whether the time intervals τ_u, τ_r and $\Delta\tau$ changed over time and were correlated with pH, base deficit and level of lactates, along the entire SEVERE phase. According to the stimulation protocol, blood samples were collected every 20 min up to the termination of the study. The same morphological parameters of Equations (1) and (2) were therefore estimated on all 20 min windows preceding each blood sample. Two correlation analyses were thus performed. First, we computed the correlation coefficients between τ_u, τ_r and $\Delta\tau$, and blood sample time points. Second, we determined the correlation between τ_u, τ_r and $\Delta\tau$ with the biomarkers (pH, BE and lactate). To compensate for the fact that biomarkers' values changed over time according to the stimulation protocol, partial correlation coefficients were computed, by accounting for the progress of time.

2.5. PRSA, AC, DC, DR, and DA

A complete description of the PRSA algorithm can be found in Bauer et al. (5) and Rivolta et al. (12). The algorithm is divided into two steps. First, anchor points are identified on the time series $x[k]$. Each time index k that satisfies the condition

$$\frac{1}{T}\sum_{i=0}^{T-1} x[k+i] > \frac{1}{T}\sum_{i=1}^{T} x[k-i], \quad (3)$$

is inserted in the DC anchors' point list (for AC, the inequality sign must be flipped). The parameter T refers to the number of FRR intervals involved in the identification of anchor points. Second, all the windows of $2L$ elements centered on each anchor point are aligned (anchor points are located at the $L+1$ sample) and then averaged. The parameter L defines the slowest FRR oscillation detectable by the algorithm (5). Such series of $2L$ averaged elements is the PRSA series.

From the PRSA series, AC and DC are then derived with

$$\text{DC (or AC)} = \frac{\sum_{i=1}^{s} \text{PRSA}[L+i] - \text{PRSA}[L-i+1]}{2s}. \quad (4)$$

DR is instead defined as the sum of DC and AC (note that AC is a negative quantity for RR series) (12).

We quantified AC, DC and DR for multiple values of T ($s = T$ and $L = 50$). A correlation analysis was performed to assess which range of T mostly correlated with the time constants derived from the piecewise linear models. In particular, we computed the correlation between τ_u and DC, τ_r and AC, and $\Delta\tau$ and DR while varying the T value between 1 and 50.

Given the fact that: i) AC, DC and DR depend on the power of the series (12); ii) a difference in the power of FRR signals was previously observed between the normoxic and hypoxic fetuses in this dataset (12); and iii) a high variability in the deceleration depth a was observed (see section 3.1), we computed the partial correlation coefficients, while compensating for the amplitude a and baseline b. In this way, correlations were insensitive to linear relations of such quantities.

Finally, we derived the relationship between the PRSA-based parameters and DA. To do so, we first computed the DA by means of the area of the trapezoid (after converting the FRR into the FHR) for both normoxic and chronically hypoxic fetuses. Then, we compared the values obtained and discussed the relationship found with the PRSA-based parameters.

2.6. Statistical Analysis

Results are reported as mean ± standard deviation and quantities were compared between the normoxic and hypoxic fetuses using a student t-test. Correlations and partial correlations were computed using the Pearson's correlation coefficient. Considering the low sample size, hypothesis tests and correlation coefficients were considered statistically significant when $p < 0.1$ (we also specify when $p < 0.05$).

3. RESULTS

3.1. Comparison of Morphological Parameters Between Normoxic and Hypoxic Sheep Fetuses

The morphological parameters obtained after model fitting were compared between the normoxic and hypoxic sheep fetuses. We obtained a model fitting achieving R^2 values of 0.8 ± 0.1. **Figure 2** reports the scatter plots for all pairs of morphological parameters for both animal models. A clear difference was found for τ_u (normoxic vs. hypoxic; 24.8 ± 9.4 vs. 39.8 ± 9.7 s; $p < 0.05$), no difference for τ_r (11.6 ± 4.8 vs. 16.0 ± 3.9 s; $p > 0.1$), a difference in FRR change a (268.1 ± 109.5 vs. 373.0 ± 46.0 ms; $p < 0.1$), and no difference for the baseline b (357.0 ± 34.1 vs. 372.6 ± 23.6 ms; $p > 0.1$). $\Delta\tau$ was found different between normoxic and hypoxic fetuses (13.2 ± 6.9 vs. 23.9 ± 7.5 s; $p < 0.05$).

3.2. Time Progression of FHR Deceleration Morphology and Correlation With Biomarkers

A correlation analysis was performed to assess the time progression of the morphological parameters over the entire duration of the SEVERE phase. We found no correlation between the morphological parameters and time for both animal models when considered together (correlations between time vs. τ_u, τ_r, and $\Delta\tau$ were -0.2, -0.1, and -0.3, respectively; $p > 0.1$) or separated (normoxic fetuses: -0.1, 0.1, and -0.1 and hypoxic fetuses: 0.1, 0.3, and -0.1; $p > 0.1$). Non-significant correlations were likely due to the limited sample size. In fact, a similar analysis performed on the biomarkers resulted in a moderate (significant) correlation with time (pH vs. time: -0.5, $p < 0.05$; BE vs. time: -0.6, $p < 0.05$; lactate vs. time: 0.34, $p < 0.1$).

Consequently, when correlation was assessed between morphological parameters and biomarkers, partial correlation coefficients were calculated to account for this possible time variation. Partial correlations were found statistically significant for pH vs. τ_u (-0.5; $p < 0.05$) and pH vs. $\Delta\tau$ (-0.6; $p < 0.05$), and for BE vs. τ_u (-0.6; $p < 0.05$) and BE vs. $\Delta\tau$ (-0.7; $p < 0.05$), whereas no significant correlation was found for lactate. Such significant correlations were mostly due to the normoxic dataset (pH vs. τ_u: -0.5, pH vs. $\Delta\tau$: -0.6, BE vs. τ_u: -0.6, BE vs. $\Delta\tau$: -0.7; $p < 0.05$). Indeed, pH and BE were found correlated with the morphological parameters only for the normoxic data ($p < 0.05$). On the other hand, lactate was found correlated with only the morphological parameters of the hypoxic fetuses (coefficients for τ_u and $\Delta\tau$ were -0.8 and -0.8; $p < 0.05$).

3.3. Correlation Analysis Between PRSA and Morphological Parameters, and DA Computation

Partial correlation coefficients were calculated between the PRSA and morphological parameters on both animal models. Partial correlation coefficients were computed to compensate for the a and b values. An example of FRR after this compensation, for both animal models, is shown in **Figure 3**. A slower FHR adaptation, as response to UCO, becomes clearly visible for the hypoxic fetus. A wide range of statistically significant correlations was found for DC vs. τ_u, AC vs. τ_u and DR vs. $\Delta\tau$, reaching values larger than -0.9 ($p < 0.1$) for $T < 20$. In particular, the normoxic fetuses showed correlations between $T = 6$ and $T = 18$ (**Figure 4A**), whereas the hypoxic ones between $T = 1$ and $T = 13$ (**Figure 4B**).

For both animal models, DA was quantified and a comparison between the DA of the two populations was performed. DA was not found significantly different between normoxic and hypoxic fetuses in our data (62.9 ± 19.5 vs. 64.4 ± 6.6 beats; $p > 0.1$).

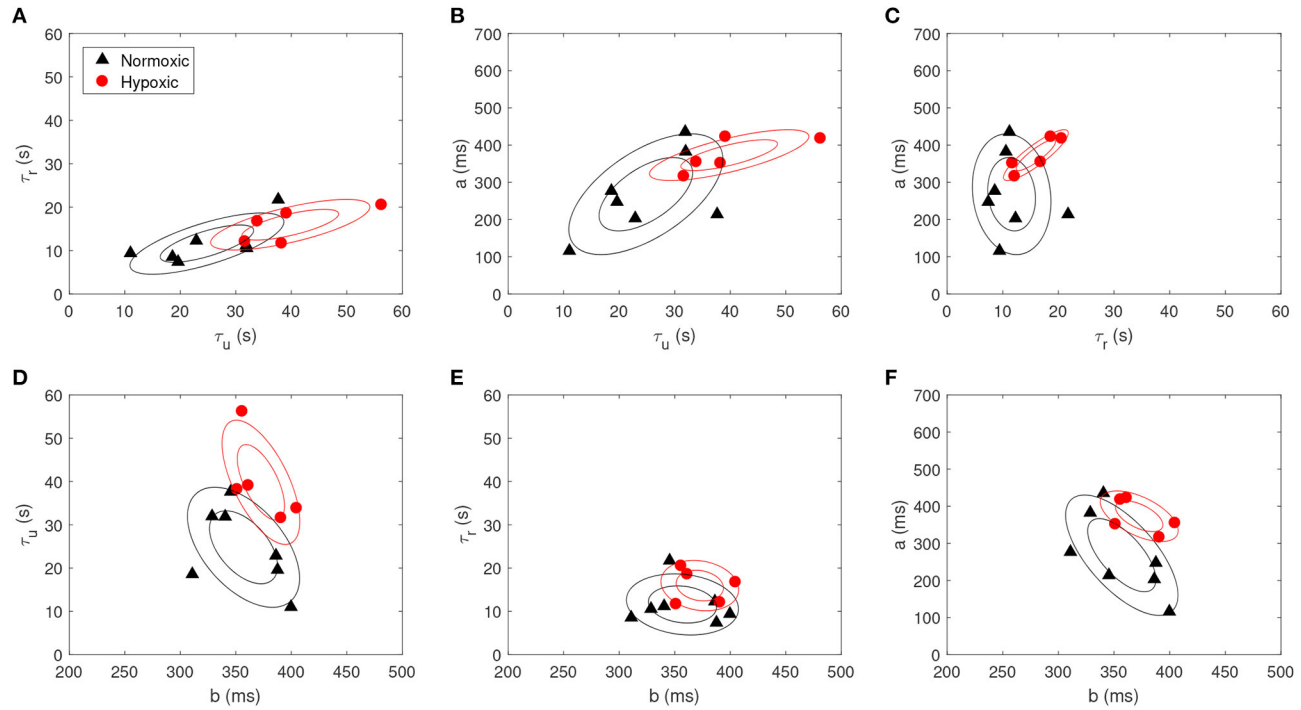

FIGURE 2 | Scatter plot for each pair of the morphological parameters, *i.e.*, **(A)** τ_u vs. τ_r, **(B)** τ_u vs. a, **(C)** τ_r vs. a, **(D)** b vs. τ_u, **(E)** b vs. τ_r and **(F)** b vs. a, for both normoxic (black triangle) and hypoxic fetuses (red circle). Contour plots are also reported (prepared with an assumption of Gaussianity, made for visualization purposes).

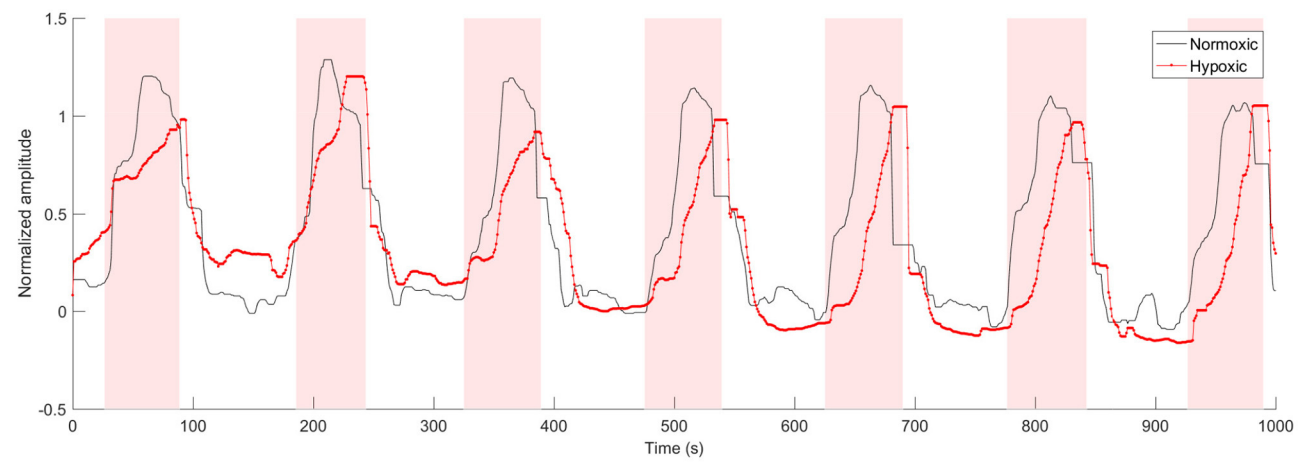

FIGURE 3 | Example of normalized FRR signals from one normoxic and one hypoxic fetus (a median filter was applied to the signals to enhance the trend). The shaded area corresponds to UCOs.

4. DISCUSSION

4.1. Morphological Differences in FHR Decelerations

When comparing the morphological differences of FHR decelerations, the significant differences in τ_u, $\Delta\tau$ and a between normoxic and chronically hypoxic fetuses suggest a different FHR response to the ANS stimulation caused by UCOs in hypoxic fetus. These findings are in line with other reports of different response to external hypoxic stimulation, as in our case by UCOs, in normoxic vs. already hypoxic fetuses (16). The reduction of FHR in the presence of hypoxic stimuli in a normoxic fetus represents a protective mechanism while it reduces the oxygen consumption via reduced myocardial work (17). In the presence of already established acidemia, the ANS modulation changes (18, 19), and some of the adaptive mechanisms, such

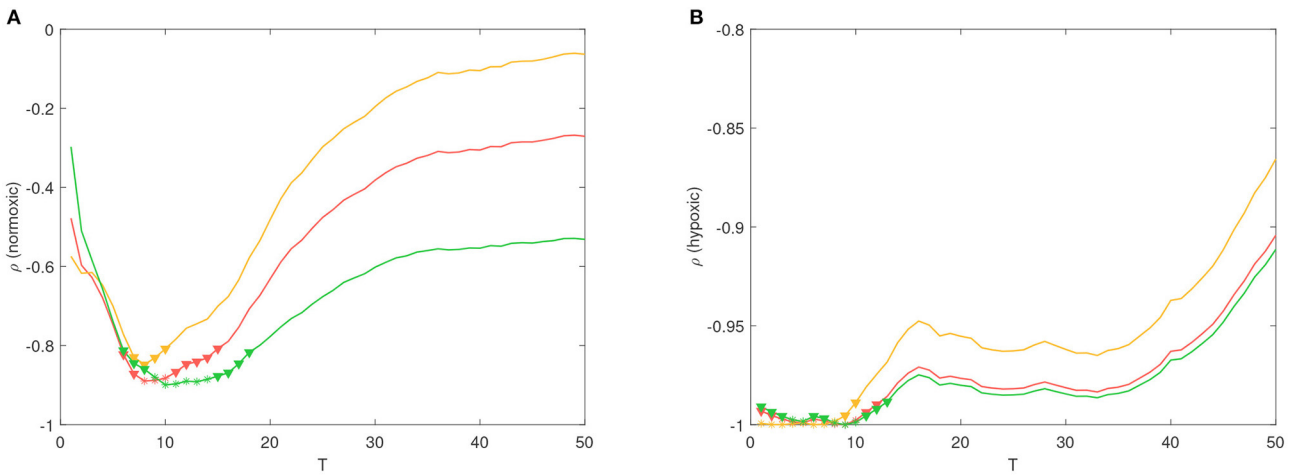

FIGURE 4 | Partial correlation coefficient ρ between τ_u vs. DC (red line), τ_r vs. AC (yellow line) and $\Delta\tau$ vs. DR (green line), for different values of T. Coefficients were computed from the **(A)** normoxic and **(B)** hypoxic fetuses. It is worth noting that the number of data points available for the computation were 7 and 5 at each T-value. $*$ ($p < 0.05$) and $\triangledown$ ($p < 0.1$) refer to statistically significant correlations.

as chemoreceptor-mediated circulatory adaptation, might be altered due to progressive tissue damage, including brain damage. In this view, the fact that the correlation between biomarkers (pH and BE) and time was driven mainly by the normoxic fetuses should not be surprising. Both pH and BE are strong stimulators of chemoreceptors in normoxic fetus (20). Thus, we speculate that the absence of the correlation between pH and BE in hypoxic fetus might represent a sign of already altered ANS and chemoreceptor function which are no longer able to respond and contribute to maintaining the balance of the aerobic metabolism. Moreover, large observational studies shows that the risk of neonatal morbidity only increases with very low pH values ($<$ 7.00) (21, 22). Certainly, it has to be acknowledged that the number of available measurements was different for normoxic and hypoxic fetuses (21 vs. 9), resulting in a much shorter SEVERE phase for the hypoxic ones. On the other hand, we found that the lactate correlated only with the morphological parameters in the hypoxic fetuses. A possible explanation could be due to the fact that the lactate represents the end product of anaerobic glucose metabolism, reflecting, thus, the metabolic acidosis (23). Indeed, it has been suggested that lactate concentrations, on fetal scalp and umbilical artery at birth, might be a better predictor of poor neonatal outcome than pH (24) and even if its use has not been universally adopted, lactate monitoring is recommended in several national and international guidelines (23).

4.2. Correlation Between Morphological Parameters and AC, DC and DR

Morphological parameters τ_u, τ_r and $\Delta\tau$ were found to correlate with DC, AC and DR for T values less than 20. In our previous study on the same animal model, we found that DR achieved the highest discriminatory power in distinguishing between normoxic and hypoxic sheep fetuses during the SEVERE phase of the protocol, with T ranging between 5 and 9 (12). Such high discriminatory power was likely related to the larger $\Delta\tau$ measured on hypoxic fetuses in this study. In other words, we found that the ANS regulation of already hypoxic fetuses during labor affects the deceleration morphology (particularly $\Delta\tau$), thus further highlighting that the presence of asymmetric trends in the series is relevant for risk stratification.

Correlations were not statistically significant for the entire range of T values considered. This was an expected result because the T value acts as frequency selector, specifically as a band-pass filter (25), whose frequency band shrinks when T increases. Although there is no clear evidence about the optimal T value for the detection of already hypoxic fetuses, previous studies employed effectively, for the detection of intra-uterine growth restriction (IUGR) during antepartum fetal monitoring, values of T corresponding to the range 2.5 s to 10 s (6, 9, 26). On the other hand, fetal acidemia occurring during labor seems better detected at lower time scales between 0.5 and 1.25 s (12, 27), thus suggesting a different mechanism for healthy fetuses during acute stress.

The PRSA series is also amplitude-dependent. In our previous study, we found a perfect linear relationship between the standard deviation of the series and the PRSA parameters (12) for Gaussian processes during stationary condition (*e.g.*, a situation likely occurring during antepartum fetal monitoring). A similar relation is expected for other indices of variability, such as the short-term variation (STV). In fact, considering results obtained during fetal monitoring of IUGR fetuses, Huhn et al. (26) and Graatsma et al. (28) found a correlation of about 0.7 between STV and AC for IUGR fetuses antepartum. On the other hand, given the non-stationary nature of FHR series during labor and the fact that the PRSA algorithm is applied to the entire recording, the relationship between STV and PRSA may break, as supported by the study of Georgieva et al. (27) who reported a significant correlation of about 0.3 during labor. It sounds therefore reasonable that the long-term variability of FHR series

may better correlate with PRSA parameters during labor. In fact, the deceleration depth a affects the variability of the entire FHR series (the standard deviation of the FHR series is proportional to a), and in turn, it affects the values of AC, DC and DR. However, while a was found larger in the hypoxic dataset, it is still unclear whether such morphological parameter would turn out to be important for risk stratification or needs to be considered as a confounding factor.

Assessing whether AC, DC and DR reflect changes in the acid/base balance is challenging. On one hand, for the normoxic fetuses, we found that i) AC/DC varied across the phases of the stimulation protocol and were correlated with changes of the biomarkers (10), ii) morphological parameters τ_u and $\Delta\tau$ were correlated to pH and BE in the SEVERE phase, and iii) correlations between morphological parameters τ_u vs. DC, τ_r vs. AC, and $\Delta\tau$ vs. DR were observed in the SEVERE phase. In this normoxic condition, the acid/base balance of the healthy fetus is continuously maintained by the placenta and its gas exchanges. The establishment of hypoxia, and a subsequent metabolic acidosis, depends both on the extent, duration and repetitiveness of the events responsible for a reduced supply of fetal oxygen, and on the metabolic reserves of the fetus at the onset of labor (29). These findings suggest that the advancing labor, along with the increased intensity and frequency of uterine contractions, causes fetal ANS adaptation that directly affects the acid/base balance and produces modifications of the deceleration morphology. The latter is thus captured by the PRSA algorithm. Our speculation is corroborated by the superior performance of PRSA-based parameters to detect acidemia at birth with respect to STV (27). On the other hand, for chronically hypoxic fetuses, we found that only lactate correlated with τ_u and $\Delta\tau$ during the SEVERE phase. It is difficult to speculate about additional relationships between PRSA-based parameters and the acid/base balance during labor for these fetuses. In fact, to the best of our knowledge, no studies quantified the correlation between PRSA-based parameters and the biomarkers in animal models of chronic hypoxia. However, several are the clinical studies suggesting that the PRSA-based parameters may differentiate between healthy and IUGR fetuses (whose our chronically hypoxic sheep are a model of) at different gestational age (6, 8) or detecting the short-term outcome for these compromised fetuses (9). Thus, confirming that the PRSA algorithm detects a deterioration of the acid/base balance of chronically hypoxic fetuses, rather than other factors, is at this point difficult to claim. We leave the quantification of the correlation between biomarkers and PRSA-based parameters on our chronically hypoxic sheep model for future works.

4.3. Comparison With Deceleration Area

The results reported so far are in line with other attempts of "capturing" the deceleration morphology for risk stratification. For example, the well-known DA quantifies the severity of the deceleration taking into account both its depth and duration (the number of "missed" beats due to the deceleration). In the work of Cahill et al. (30), DA was quantified by approximating the deceleration by a triangle having a base as long as its duration and height corresponding to the depth, and then computing its area. DA was found to perform well for detecting fetal acidemia (AUC $=$ 0.76). Using the morphological parameters, which are depicted in **Figure 1**, DA is then given by

$$\mathrm{DA} = \frac{1000}{2} \times \frac{d_{\mathrm{UCO}} + \tau_r}{60} \times \left(\frac{60}{b} - \frac{60}{b+a} \right) \quad \text{(beats)}, \quad (5)$$

where $d_{\mathrm{UCO}} = 60$ s is the duration of the UCO, and $1,000$ is the conversion factor from milliseconds to seconds. In our case, given the severity of the stimulation protocol, UCOs caused FHR responses which were better approximated by a trapezoidal model than a triangle. Thus, it was possible to calculate DA using also the following formula

$$\mathrm{DA} = \frac{1000}{2} \times \frac{2d_{\mathrm{UCO}} - \Delta\tau}{60} \times \left(\frac{60}{b} - \frac{60}{b+a} \right) \quad \text{(beats)}. \quad (6)$$

From these formulas, the relationship between DA and the morphological parameters, which we studied in this paper, is evident. The second formula also points out a link between DA and the asymmetry value $\Delta\tau = \tau_u - \tau_r$, which is not captured in the triangular approximation usually employed. The relationship with $\Delta\tau$ hints the importance of looking at asymmetric trends in the series present during labor. We leave this point to future investigations.

In our study, however, DA was not found significantly different between normoxic and hypoxic fetuses in our data (62.9 ± 19.5 vs. 64.4 ± 6.6 beats; $p > 0.1$). A possible explanation could be the fact that $\Delta\tau$ and a were higher for the hypoxic fetuses, and such quantities correlated with DA in opposite directions, thus making DA values indistinguishable due to balancing effects and (likely) the limited sample size. Another possible reason is that chronic hypoxia (in this study) and acidemia at birth (30) might trigger different regulatory ANS responses.

4.4. Clinical Implications

Over the past 50 years, the cardiotocography has become the most frequently used method for intrapartum surveillance of fetal wellbeing. Intrapartum FHR monitoring, in particular as part of the CTG and computerized systems such as those developed by Dawes and Redman, has reduced perinatal mortality (31). In 1991, Dawes et al. (32) reported that low short-term variation was associated with adverse perinatal outcome, however its application required antepartum monitoring to identify at-risk fetuses which to-date has not become the standard of care. Recent developments in the remote transabdominal ECG technology are enabling the antepartum monitoring, so more progress is to be expected in incorporating various computerized FHR algorithms into the standard of care in the coming years (33). In the meantime, the intrapartum CTG as currently practiced in over 90% of the delivering hospitals remains to be characterized by high intra and inter-observer interpretative variability, and by low specificity in identifying fetal acidosis and fetal acidosis at birth. Its poor performance is compounded by the fact that fetal acidosis itself is a poor predictor of perinatal brain injury (34). Thus, CTG does not reduce perinatal morbidity and mortality, but it does increase the rate of operative deliveries (35, 36).

Indeed, there is still debate regarding CTG application during labor (37). Randomized controlled trials failed to demonstrate improved outcomes with the use of CTG during labor (38, 39). Moreover, recent evidence suggests that the type of deceleration [defined by NICHD FHR parameters criteria (40)] does not predict low pH at birth (30), and that the deceleration area is the most discriminative in the identification of fetal acidemia (41). Thus, the authors concluded that there is no need to determine the type of deceleration at all, which is prone to high interobserver variation. However, this statement is not shared by other authors (42), arguing that the success of CTG depends on team effort by the obstetricians with in-depth practical knowledge.

Since singular parameters generally have been shown to have a poor predictive value, a multiparametric approach to FHR patterns is also emphasized. Recently, Eden et al. (43) developed a new promising multiparametric metric, the fetal reserve index, that showed a reduction in emergency operative deliveries and in adverse fetal/neonatal outcomes (44, 45). Building upon the notion of multiparametric approaches and the focus on fetal outcomes instead of poorly correlated outcome metrics such as fetal acidemia and leveraging deep learning, recent advances have been made in detecting preventable fetal distress and brain injury from multiple CTG patterns (34). Overall, there is evidence to suggest the new metrics and types of analysis of FHR might improve what, at the moment, represents the standard clinical practice. Based on our result, AC, DC and DR, computed through PRSA, are linked to the deceleration morphology itself. Thus, they may lead to a better categorization of the deceleration type. Moreover, the deceleration morphologies have been found to differ between normoxic and hypoxic sheep fetuses during UCO, thus, motivating future investigation on the translational potential of this methodology. Clinical cohorts are needed to validate these findings and evaluate the clinical performance of these new metrics in identifying compromised fetuses during labor.

4.5. Limitations of the Study

First, the sample size is small, dictated by the complexity of the animal model. Second, UCOs implemented in the experiments do not necessarily generalize to human labor, where the contractions are not equally regular nor are they all producing a complete occlusion of the umbilical cord. Third, sheep were analyzed during complete UCOs. However, changes in ANS activity in response to UCOs also occur earlier in time, when the UCOs are less severe or the recovery time between the UCOs is longer, and may reflect differences in the chronically hypoxic fetuses compared to the normoxic ones. Identifying these potentially earlier differences will be the subject of future studies. Fourth, all sheep fetuses displayed an individual pattern of pathological hypotensive responses to UCOs with regard to the timing of its emergence, with hypotensive responses to FHR decelerations showing well ahead of the severe UCOs in some instances (46, 47). As our present study focused on the differences between the hypoxic and normoxic fetuses in the severe stage of UCOs, it did not investigate the relationship between the PRSA-based metrics and the timing of the onset of pathological hypotension. We leave this to future work.

4.6. Conclusions

Our study motivates further investigations on PRSA-related quantities to determine their potential advantage for risk stratification. It might also open interesting scenarios for interpreting PRSA-based results and improving FHR monitoring. The evaluation of the performance of these new metrics in identifying compromised fetuses during labor is still underway.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The dataset can be made available upon request. Requests to access these datasets should be directed to Martin G. Frasch, mfrasch@uw.edu.

ETHICS STATEMENT

The animal study was reviewed and approved by University of Western Ontario Council on Animal Care.

AUTHOR CONTRIBUTIONS

MR designed and implemented the analyses and drafted the manuscript. MF collected the data. MR and RS optimized the proposed mathematical framework. MB, TS, and MF were involved in the clinical interpretation of the results. All authors read, revised, and approved the final manuscript.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Bryan Richardson and his Perinatal Research Lab at the University of Western Ontario for the original design of the animal experiments that enabled the acquisition of the dataset underlying the present study. The preprint version of this manuscript is available on biorxiv.org (48).

REFERENCES

1. Divon MY, Muskat Y, Platt LD, Paldi E. Increased beat-to-beat variability during uterine contractions: a common association in uncomplicated labor. *Am J Obstet Gynecol*. (1984) 149:893–6. doi: 10.1016/0002-9378(84)90611-2
2. Bennet L, Gunn AJ. The fetal heart rate response to hypoxia: insights from animal models. *Clin Perinatol*. (2009) 36:655–72. doi: 10.1016/j.clp.2009.06.009
3. Jonsson M, Agren J, Nordén-Lindeberg S, Ohlin A, Hanson U. Neonatal encephalopathy and the association to asphyxia in labor. *Am J Obstet Gynecol*. (2014) 211:667.e1-8. doi: 10.1016/j.ajog.2014.06.027

4. Alfirevic Z, Devane D, Gyte GM, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev*. (2017) 2:CD006066. doi: 10.1002/14651858.CD006066.pub3
5. Bauer A, Kantelhardt JW, Bunde A, Barthel P, Schneider R, Malik M, et al. Phase-rectified signal averaging detects quasi-periodicities in non-stationary data. *J Phys A*. (2006) 364:423–34. doi: 10.1016/j.physa.2005.08.080
6. Stampalija T, Casati D, Montico M, Sassi R, Rivolta MW, Maggi V, et al. Parameters influence on acceleration and deceleration capacity based on trans-abdominal ECG in early fetal growth restriction at different gestational age epochs. *Eur J Obstet Gynecol Reprod Biol*. (2015) 188:104–12. doi: 10.1016/j.ejogrb.2015.03.003
7. Stampalija T, Casati D, Monasta L, Sassi R, Rivolta MW, Muggiasca ML, et al. Brain sparing effect in growth-restricted fetuses is associated with decreased cardiac acceleration and deceleration capacities: a case-control study. *BJOG*. (2016) 123:1947–54. doi: 10.1111/1471-0528.13607
8. Tagliaferri S, Fanelli A, Esposito G, Esposito FG, Magenes G, Signorini MG, et al. Evaluation of the acceleration and deceleration phase-rectified slope to detect and improve IUGR clinical management. *Comput Math Methods Med*. (2015) 2015:1–9. doi: 10.1155/2015/236896
9. Lobmaier SM, van Charante NM, Ferrazzi E, et al. Phase-rectified signal averaging method to predict perinatal outcome in infants with very preterm fetal growth restriction- a secondary analysis of TRUFFLE-trial. *Am J Obstet Gynecol*. (2016) 215:630.e1–630.e7. doi: 10.1016/j.ajog.2016.06.024
10. Rivolta MW, Stampalija T, Casati D, Richardson BS, Ross MG, Frasch MG, et al. Acceleration and deceleration capacity of fetal heart rate in an *in-vivo* sheep model. *PLoS ONE*. (2014) 9:e104193. doi: 10.1371/journal.pone.0104193
11. 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) 57:771–9. doi: 10.1109/TBME.2009.2035818
12. Rivolta MW, Stampalija T, Frasch MG, Sassi R. Theoretical value of deceleration capacity points to deceleration reserve of fetal heart rate. *IEEE Trans Biomed Eng*. (2020) 67:1176–85. doi: 10.1109/TBME.2019.2932808
13. Morrison JL, Berry MJ, Botting KJ, Darby JRT, Frasch MG, Gatford KL, et al. Improving pregnancy outcomes in humans through studies in sheep. *Am J Physiol Regul Integr Comp Physiol*. (2018) 315:R1123–53. doi: 10.1152/ajpregu.00391.2017
14. Xu A, Durosier LD, Ross MG, Hammond R, Richardson BS, Frasch MG. Adaptive brain shut-down counteracts neuroinflammation in the near-term ovine fetus. *Front Neurol*. (2014) 5:110. doi: 10.3389/fneur.2014.00110
15. Durosier LD, Green G, Batkin I, Seely AJ, Ross MG, Richardson BS, et al. Sampling rate of heart rate variability impacts the ability to detect acidemia in ovine fetuses near-term. *Front Pediatr*. (2014) 2:38. doi: 10.3389/fped.2014.00038
16. Parer J, Dijkstra H, Harris J, Krueger T, Reuss M. Increased fetal heart rate variability with acute hypoxia in chronically instrumented sheep. *Eur J Obstet Gynaecol Reprod Biol*. (1980) 10:393–9. doi: 10.1016/0028-2243(80)90025-8
17. Fletcher AJW, Gardner DS, Edwards CMB, Fowden AL, Giussani DA. Development of the ovine fetal cardiovascular defense to hypoxemia towards full term. *Am J Physiol Heart Circ Physiol*. (2006) 291:H3023–34. doi: 10.1152/ajpheart.00504.2006
18. Murotsuki J, Bocking AD, Gagnon R. Fetal heart rate patterns in growth restricted fetal sheep induced by chronic fetal placental embolization. *Am J Obstet Gynecol*. (1997) 176:282–90. doi: 10.1016/S0002-9378(97)70486-1
19. Siira S, Ojala T, Vahlberg T, Rosén K, Ekholm E. Do spectral bands of fetal heart rate variability associate with concomitant fetal scalp pH? *Early Hum Dev*. (2013) 89:739–42. doi: 10.1016/j.earlhumdev.2013.05.007
20. Itskovitz J, Rudolph AM. Denervation of arterial chemoreceptors and baroreceptors in fetal lambs in utero. *Am J Physiol Heart Circ Physiol*. (1982) 242:H916–20. doi: 10.1152/ajpheart.1982.242.5.H916
21. Yeh P, Emary K, Impey L. The relationship between umbilical cord arterial pH and serious adverse neonatal outcome: analysis of 51,519 consecutive validated samples. *BJOG*. (2012) 119:824–31. doi: 10.1111/j.1471-0528.2012.03335.x
22. Mittendorf R, Won SY, Gianopoulos JG, Pryde PG, Roizen N. Relationships between umbilical cord arterial blood pH levels at delivery and bayley psychomotor development index scores in early childhood. *Perinat Med*. (2008) 36:335–40. doi: 10.1515/JPM.2008.043
23. Cummins G, Kremer J, Bernassau A, Brown A, Bridle H, Schulze H, et al. Sensors for fetal hypoxia and metabolic acidosis: a review. *Sensors*. (2018) 18:2648. doi: 10.3390/s18082648
24. Kruger K, Hallberg B, Blennow M, Kublickas M, Westgren M. Predictive value of fetal scalp blood lactate concentration and pH as markers of neurologic disability. *Am J Obstet Gynecol*. (1999) 181:1072–8. doi: 10.1016/S0002-9378(99)70083-9
25. Sassi R, Stampalija T, Casati D, Ferrazzi E, Bauer A, Rivolta MW. A methodological assessment of phase-rectified signal averaging through simulated beat-to-beat interval time series. *Comput Cardiol*. (2014) 41:601–4. Available online at: https://www.cinc.org/archives/2014/pdf/0601.pdf
26. Huhn EA, Lobmaier S, Fischer T, Schneider R, Bauer A, Schneider KT, et al. New computerized fetal heart rate analysis for surveillance of intrauterine growth restriction. *Prenat Diagn*. (2011) 31:509–14. doi: 10.1002/pd.2728
27. Georgieva A, Papageorghiou AT, Payne SJ, Moulden M, Redman CW. Phase-rectified signal averaging for intrapartum electronic fetal heart rate monitoring is related to acidaemia at birth. *BJOG*. (2014) 121:889–94. doi: 10.1111/1471-0528.12568
28. Graatsma EM, Mulder EJH, Vasak B, Lobmaier SM, Pildner von Steinburg S, Schneider KTM, et al. Average acceleration and deceleration capacity of fetal heart rate in normal pregnancy and in pregnancies complicated by fetal growth restriction. *J Matern Fetal Neonatal Med*. (2012) 25:2517–22. doi: 10.3109/14767058.2012.704446
29. Yli BM, Kjellmer I. Pathophysiology of foetal oxygenation and cell damage during labour. *Best Pract Res Clin Obstet Gynaecol*. (2016) 30:9–21. doi: 10.1016/j.bpobgyn.2015.05.004
30. Cahill AG, Tuuli MG, Stout MJ, López 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:e1–523.e12. doi: 10.1016/j.ajog.2018.01.026
31. Redman C. 45 years of fetal heart rate monitoring in BJOG. *BJOG*. (2015) 122:536. doi: 10.1111/1471-0528.13101
32. Dawes GS, Moulden M, Redman CW. The advantages of computerized fetal heart rate analysis. *J Perinat Med*. (1991) 19:39–45. doi: 10.1515/jpme.1991.19.1-2.39
33. Mhajna M, Schwartz N, Levit-Rosen L, Warsof S, Lipschuetz M, Jakobs M, et al. Wireless, remote solution for home fetal and maternal heart rate monitoring. *Am J Obstet Gynecol MFM*. (2020) 2:100101. doi: 10.1016/j.ajogmf.2020.100101
34. Frasch MG, Strong S, Nilosek D, Leaverton J, Schiffrin BS. Detection of preventable fetal distress during labor from scanned cardiotocogram tracings using deep learning. *Front Pediatr*. (2021).
35. 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) 216:163.e1–163.e6. doi: 10.1016/j.ajog.2016.10.009
36. Johnson GJ, Salmanian B, Denning SG, Belfort MA, Sundgren NC, Clark SL. Relationship between umbilical cord gas values and neonatal outcomes: implications for electronic fetal heart rate monitoring. *Obstet Gynecol*. (2021) 138:366–73. doi: 10.1097/AOG.0000000000004515
37. Frasch MG, Boylan GB, Wu HT, Devane D. Commentary: computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Front Physiol*. (2017) 8:721. doi: 10.3389/fphys.2017.00721
38. Vintzileos AM, Antsaklis A, Varvarigos I, Papas C, Sofatzis I, Montgomery JT. A randomized trial of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation. *Obstet Gynecol*. (1993) 81:899–907.
39. The-INFANT-Collaborative-Group. Computerised interpretation of the fetal heart rate during labour: a randomised controlled trial (INFANT). *Lancet*. (2017) 389:1719–29. doi: 10.1016/S0140-6736(17)30568-8
40. ACOG. ACOG Practice Bulletin No. 106: Intrapartum fetal heart rate monitoring: nomenclature, interpretation, and general management principles. *Obstet Gynecol*. (2009) 114:192–202. doi: 10.1097/AOG.0b013e3181aef106
41. Furukawa A, Neilson D, Hamilton E. Cumulative deceleration area: a simplified predictor of metabolic acidemia. *J Matern Fetal Neonatal Med*. (2019) 30:1–8. doi: 10.1080/14767058.2019.1678130

42. Sholapurkar SL. The present and future of intrapartum computerized cardiotocography: role of pattern recognition incorporating single vs. multiple parameters. *J Matern Fetal Neonatal Med*. (2021) 17:1–7. doi: 10.1080/14767058.2021.1949453
43. Eden RD, Evans MI, Evans SM, Schifrin BS. The "Fetal reserve Index": re-engineering the interpretation and responses to fetal heart rate patterns. *Fetal Diagn Ther*. (2018) 43:90–104. doi: 10.1159/000475927
44. Eden RD, Evans MI, Britt DW, Evans SM SB. Safely lowering the emergency cesarean and operative vaginal delivery rates using the fetal reserve index. *J Matern Fetal Neonatal Med*. (2020) 33:1473–9. doi: 10.1080/14767058.2018.1519799
45. Evans MI, Eden RD, Britt DW, Evans SM SB. Re-engineering the interpretation of electronic fetal monitoring to identify reversible risk for cerebral palsy: a case control series. *J Matern Fetal Neonatal Med*. (2019) 32:2561–9. doi: 10.1080/14767058.2018.1441283
46. Gold N, Frasch MG, Herry CL, Richardson BS, Wang X. A doubly stochastic change point detection algorithm for noisy biological signals. *Front Physiol*. (2018) 8:1112. doi: 10.3389/fphys.2017.01112
47. Gold N, Herry CL, Wang X, Frasch MG. Fetal cardiovascular decompensation during labor predicted from the individual heart rate: a prospective study in fetal sheep near term and the impact of low sampling rate. *arXiv e-prints*. (2019) arXiv:1911.01304.
48. Rivolta MW, Barbieri M, Stampalija T, Sassi R, Frasch MG. Relationship between deceleration morphology and phase rectified signal averaging-based parameters during labor. *bioRxiv*. (2021) doi: 10.1101/2021.04.21.440741

Conflict of Interest: MF has a patent pending on abdominal ECG signal separation for FHR monitoring (WO2018160890).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Medicine

SYSTEMATIC REVIEW
published: 30 November 2021
doi: 10.3389/fmed.2021.661226

Non-linear Methods Predominant in Fetal Heart Rate Analysis: A Systematic Review

Maria Ribeiro[1,2]*, *João Monteiro-Santos*[3,4], *Luísa Castro*[3,4,5], *Luís Antunes*[1,2], *Cristina Costa-Santos*[3,4], *Andreia Teixeira*[3,4,6†] and *Teresa S. Henriques*[3,4†]

[1] Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal, [2] Computer Science Department, Faculty of Sciences, University of Porto, Porto, Portugal, [3] Centre for Health Technology and Services Research, Faculty of Medicine University of Porto, Porto, Portugal, [4] Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal, [5] School of Health of Polytechnic of Porto, Porto, Portugal, [6] Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal

The analysis of fetal heart rate variability has served as a scientific and diagnostic tool to quantify cardiac activity fluctuations, being good indicators of fetal well-being. Many mathematical analyses were proposed to evaluate fetal heart rate variability. We focused on non-linear analysis based on concepts of chaos, fractality, and complexity: entropies, compression, fractal analysis, and wavelets. These methods have been successfully applied in the signal processing phase and increase knowledge about cardiovascular dynamics in healthy and pathological fetuses. This review summarizes those methods and investigates how non-linear measures are related to each paper's research objectives. Of the 388 articles obtained in the PubMed/Medline database and of the 421 articles in the Web of Science database, 270 articles were included in the review after all exclusion criteria were applied. While approximate entropy is the most used method in classification papers, in signal processing, the most used non-linear method was Daubechies wavelets. The top five primary research objectives covered by the selected papers were detection of signal processing, hypoxia, maturation or gestational age, intrauterine growth restriction, and fetal distress. This review shows that non-linear indices can be used to assess numerous prenatal conditions. However, they are not yet applied in clinical practice due to some critical concerns. Some studies show that the combination of several linear and non-linear indices would be ideal for improving the analysis of the fetus's well-being. Future studies should narrow the research question so a meta-analysis could be performed, probing the indices' performance.

Keywords: fetal heart rate, non-linear methods, entropy, data compression, fractal analysis, wavelet analysis, systematic review

OPEN ACCESS

Edited by:
Patrice Abry,
École Normale Supérieure de Lyon, Université de Lyon, France

Reviewed by:
Maria G. Signorini, Politecnico di Milano, Italy
Giovanni Magenes, University of Pavia, Italy

***Correspondence:**
Maria Ribeiro
maria.r.ribeiro@inesctec.pt

†*These authors share last authorship*

Specialty section:
This article was submitted to Obstetrics and Gynecology, a section of the journal Frontiers in Medicine

Received: *30 January 2021*
Accepted: *04 November 2021*
Published: *30 November 2021*

Citation:
Ribeiro M, Monteiro-Santos J, Castro L, Antunes L, Costa-Santos C, Teixeira A and Henriques TS (2021) Non-linear Methods Predominant in Fetal Heart Rate Analysis: A Systematic Review. Front. Med. 8:661226. doi: 10.3389/fmed.2021.661226

1. INTRODUCTION

Worldwide, it is estimated that the number of fetal deaths after week 20 of gestational age is around 2.6 million per year. Although the numbers have been decreasing in the past decades, the stillbirths' rate still ranges from about 1 in 250 births in developed countries and 1 per 33 in South Asia and Sub-Saharan Africa (data from 2009), according to Cousens et al. (1).

Cardiotocography (CTG) combines fetal heart rate (fHR) measurement, obtained through a uterine contraction monitoring probe and a Doppler ultrasound probe for fHR, recorded using an abdominal pressure transducer. In developed countries, clinical decisions during labor are firmly based on fHR monitoring (2, 3), being CTG the most used tool to assess fetal well-being since the early '60s according to Spencer (4). However, the information provided by CTG is limited since a complete electrocardiogram (ECG) signal of the fetus is not available. Moreover, the CTG is highly sensitive to both fetal and maternal movement. The use of an electrode placed on the fetus's scalp is more reliable as it retrieves fetal electrocardiogram, containing not only fHR but also other crucial clinical parameters (5, 6). On the other hand, this is an invasive method only possible during labor, after the beginning of cervical dilatation and the membranes' rupture, carrying with it risks of infection (7, 8). However, other methods for fetal monitoring are used such as fetal phonocardiography (9–11), fetal echocardiography (12, 13), and fetal magnetocardiography (14, 15). Each one of the methods has its own advantages and disadvantages. For more detail on this matter, see Jaros et al. (16) and Hoyer et al. (17).

Electronic fetal monitoring came with high expectations since it offered continuous monitoring, compared to the intermittent auscultation done until then. However, a meta-analysis of large multicenter studies did not prove any significant improvement. Also, electronic fetal monitoring became the main suspect for the increased rate of cesarean sections (18). These procedures result in a slight increase in poor outcomes in low-risk pregnancies. The cesarean sections also require a longer time to heal than a vaginal birth and present increased risks, including baby breathing problems, amniotic fluid embolism, and postpartum bleeding for the mother (19). Despite the importance of the fetus and mother well-being assessment, low concordance between physicians is still present, even among experienced obstetricians, resulting in a high rate of false-positives (2, 20, 21). In daily practice, fHR is subject to the clinician visual interpretation, even when following the guidelines provided by the International Federation of Obstetrics and Gynaecology (FIGO) (22, 23), which although being associated with high sensitivity but low specificity (24), might leads to a chance of more harmful than beneficial adherence to conventional guidelines (25).

The autonomic nervous system (ANS) is involved in the control of almost every organ system, and the beat-to-beat variation of fHR reflects the influence of the fetus' ANS and its components (sympathetic and parasympathetic) and, therefore, is an indicator of fetal well-being (8). A certain level of unpredictable fetal heart rate variability (fHRV) reflects sufficient capabilities of the organism in search of optimal behavior. Reduced fHRV is linked with limited capabilities and mental disorders (26). The linear modeling approaches quantify sympathetic and parasympathetic control mechanisms and their balance by measuring spectral low and high-frequency components. However, it has been shown that not all information carried by beat-to-beat variability can be explained by these components (27). For this matter, in the past couple of decades, and with the fast development of computation, new signal processing, and pattern recognition methodologies have been developed and applied to many different fields, including the analysis of fHRV using non-linear parameters (28, 29). These approaches can reveal relevant clinical information not exposed by temporal or frequency analysis (30).

Variability and complexity are different terms. While a complex system requires variability, the other way around is not guaranteed. For example, a set of random notes in music can be interpreted as having high complexity for its non-predictability, whereas a set of consecutive notes is highly predictable, and both have high variability. Thus, complexity signals, such as those produced by self-regulatory physiological systems, present temporal or spatial structures over a varied range of scales. Because of their non-linearity and non-stationarity, conventional indicators, such as the mean and the standard deviation, do not fulfill their purpose (31). In the end, complexity is a property of any system that quantifies the amount of structured information.

Chaffin et al. (32), in 1991, were the first to use non-linear analyzes in fHR. The authors applied fractal analysis (correlation dimension) to study 12 normal fetuses' well-being in labor. Later, in 1992, Pincus and Viscarello (33) found statistically significant results using approximate entropy (ApEn) when comparing a group of acidemic fetuses with non-acidemic ones. These results supported the hypothesis that regular fHR patterns are associated with acidemia. Datian and Xuemei (34), in 1996, introduced a new wavelet analysis method used to detect fetal electrocardiogram from the abdominal signal and compared to other methods in practice. Signorini et al. (35), in 2005, applied data compression (Lempel Ziv complexity) for the first time in the fHR analysis to improve the early detection of fetal distress conditions such as intrauterine growth restriction. The same authors, also in 2007 (36), used the Lempel Ziv complexity to successfully discriminate between severe intrauterine growth restriction (IUGR) (premature birth) and non-severe IUGR (term delivery) and normal fetuses. In the subsequent year, using a compressor-based clustering algorithm called normalized compression distance (NCD), Santos et al. (37) managed to clustered abnormal and suspicious tracks, regardless of the monitoring system used. Barquero-Pérez et al. (38) also used NCD for automatic detection of perinatal hypoxia.

The main contribution of this article is to provide a systematic review of articles that apply entropy, compression, fractal, and wavelet analysis to study the dynamics of fHR and analyze the research objectives of these articles. As far as we know, there is no systematic review for this purpose in the literature.

We begin by describing the methodology used, specifying the sources of information, the eligibility criteria, the study selection, data extraction, and quality assessment in section 2. Based on the systematic review results, we describe in detail the most commonly used non-linear methods to assess the dynamics of fetal heart rate and analyze how the study of the complexity of fHR has evolved over the years (section 3). In section 4, we describe the most frequent goals in research. We analyze the evolution of the non-linear methods' applications to these objectives and probe how the research objectives are related to non-linear methods. In section 5, we reflect on some open questions regarding the application of non-linear measurements

TABLE 1 | Online queries in Pubmed and Web of Science.

Pubmed (https://www.ncbi.nlm.nih.gov/pubmed)
("Nonlinear Dynamics"[Mesh] OR "Nonlinear Dynamics"[Title/Abstract] OR Nonlinear[Title/Abstract] OR "Entropy"[Mesh] OR Entropy[Title/Abstract] OR "Data Compression"[Mesh] OR "Data Compression"[Title/Abstract] OR Compression[Title/Abstract] OR complexity[Title/Abstract] OR "fractals"[MeSH Terms] OR fractals[Title/Abstract] OR "Wavelet Analysis"[Mesh] OR "Wavelet Analyses"[Title/Abstract] OR wavelet[Title/Abstract]) AND ("Heart Rate, Fetal"[Mesh] OR "Fetal Heart Rate"[Title/Abstract] OR "foetal heart rate"[Title/Abstract])
Web of Science (https://www.webofknowledge.com)
(TS=("Nonlinear Dynamics" OR nonlinear OR entropy OR Compression OR complexity OR fractal OR wavelet) OR TI=("Nonlinear Dynamics" OR nonlinear OR entropy OR Compression OR complexity OR fractal OR wavelet)) AND (TI=("Heart Rate, Fetal") OR TI=("Fetal Heart Rate") OR TI=("foetal heart rate") OR TS=("fetal heart rate") OR TS=("foetal heart rate"))

to fHR dynamics. We finish this paper with the main conclusions in section 6.

2. SYSTEMATIC REVIEW METHODS

This systematic review focused on original papers that include non-linear analysis, such as complexity measures, fractal approaches, and wavelets, of human fetal heart rate during ante and intrapartum. The online search was performed on Medline, through PubMed, and the Web of Science databases, searching all the papers published until the 4th of October 2020. The following terms were used as descriptors/Mesh: "non-linear dynamics," "entropy," "data compression," "complexity," "fractals," "wavelets," "fetal heart rate," "foetal heart rate." The queries used in each database can be found in **Table 1**. This study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement (39). The review protocol was not registered prospectively.

Inclusion criteria for selecting studies were the following: observational or experimental papers presenting complexity analysis of fetal heart rate; abstract found online; reported original research in peer-reviewed journals; at least one measure from the following was used in the analysis (entropy, compression, fractal, or wavelet). Papers using non-human fetal heart rate analysis, papers without an English version, reviews, case studies, dissertations, and thesis were excluded (see **Figure 1**).

All authors were involved in the selection of studies, data extraction, and quality assessment. Two authors independently assessed each title and abstract found in the databases. The full texts of potentially relevant studies have been retrieved and revised in depth. Disagreements between reviewers were resolved by consensus or by the decision of a third independent reviewer. For each article, the following data were collected: year of publication, study design, objective, sample size, measure(s) used to analyze fHR, and conclusions. Both reviewers made sure that all included papers met the criteria defined in the first stage.

A total of 603 abstracts were assessed, 368 of which retained for full-text screening. Two hundred and seventy papers were then included in the review after meeting all the criteria. **Figure 1** contains the PRISMA flow diagram for study selection, including reasons for exclusion. The most used non-linear analysis measures to study the dynamics of fHR obtained in the systematic review are also represented in **Figure 1**.

3. NON-LINEAR METHODS

Although linear indices have been extensively used in fetal monitoring for the past decades, it is established that biological systems are more complex than they appear. Non-linear measures based on concepts of chaos, fractality, and complexity have gained space and demonstrated promising results in the analysis of fetal well-being and the prediction of pathologies. The application of non-linear measures to study the dynamics of fHR has increased over the years. The non-linear methods covered by this review are entropy, compression, fractal analysis, and wavelet analysis. The results show that entropy is the most applied measure in fetal heart rate, followed by fractal analysis, wavelet analysis, and the least applied is the compression (see **Figures 1**, **2**). Although the application of entropy methods stands out, we can see that compression and wavelet analysis methods have been increasingly used in recent years (see **Figure 2**).

In the following sections, the most applied non-linear methods are described. In our systematic review, other non-linear methods were found, such as, Poincaré plot (in 18 papers), symbolic dynamics (in 12 papers), phase rectified signal average (in 10 papers), Lyapunov exponents (in 6 papers), and recurrence plot analysis (in 6 papers). However, due to the reduced number of uses, they were not described in detail. For this review, we decided to describe only the measures most applied to fHR.

3.1. Entropy

According to Shannon (40), the information within a signal can be quantified with absolute precision as the amount of unexpected data in the message (defined as entropy). Entropy, a probabilistic complexity measure used to quantify a time series's irregularity, has been widely used in physiological signal analysis. The number of papers that applied each entropy measure per year is shown in **Figure 2**. The entropy measures that were applied to at least 15 articles were: Shannon entropy (SE), approximate entropy (ApEn), sample entropy (SampEn), and multiscale entropy (MSE).

From all 270 papers included in this review, 149 (55.2%) papers applied entropy: 16 (5.9%) show results with Shannon entropy (SE), 82 (30.4%) used SampEn, 101 (37.4%) used ApEn, and 30 (11.1%) used MSE (see **Figure 1**). **Figure 3** shows the number of papers that applied measures of the entropy by year. ApEn is the most applied measure. However, in recent years the employment of SampEn and ApEn is similar.

In the literature, we found other entropy measures that appeared in less than 15 articles, such as, permutation entropy (41–45), Rnyi entropy (46–48), Kullback-Leibler entropy (41,

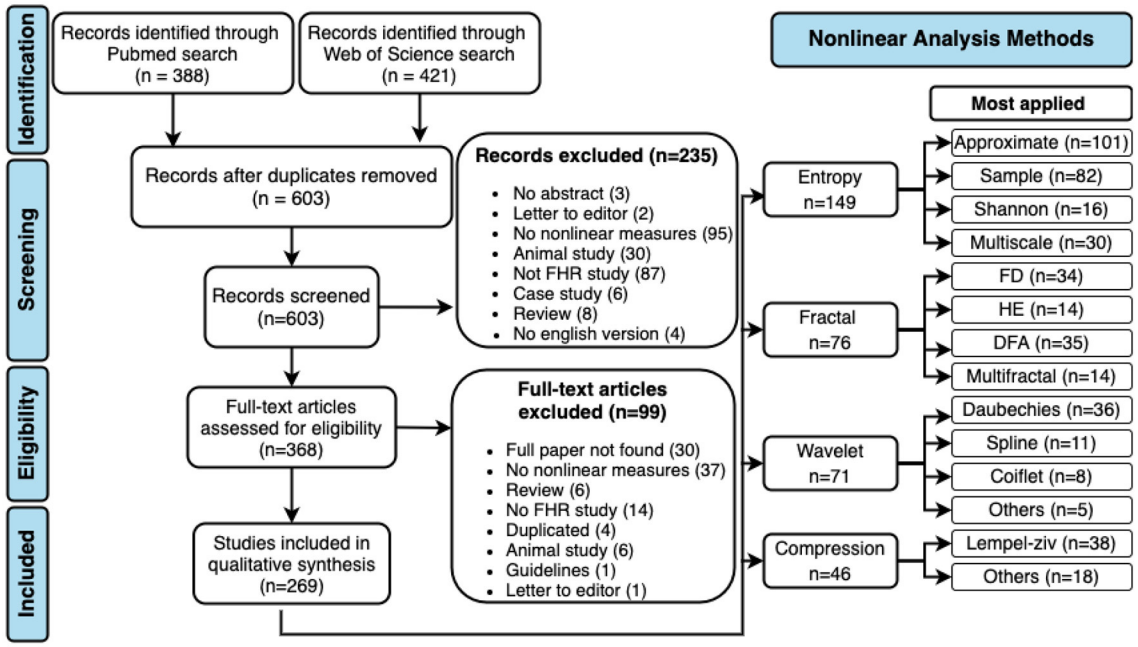

FIGURE 1 | PRISMA flow diagram and non-linear methods most applied in fHR. DFA, detrended fluctuation analysis; FD, fractal dimension; HE, Hurst exponent.

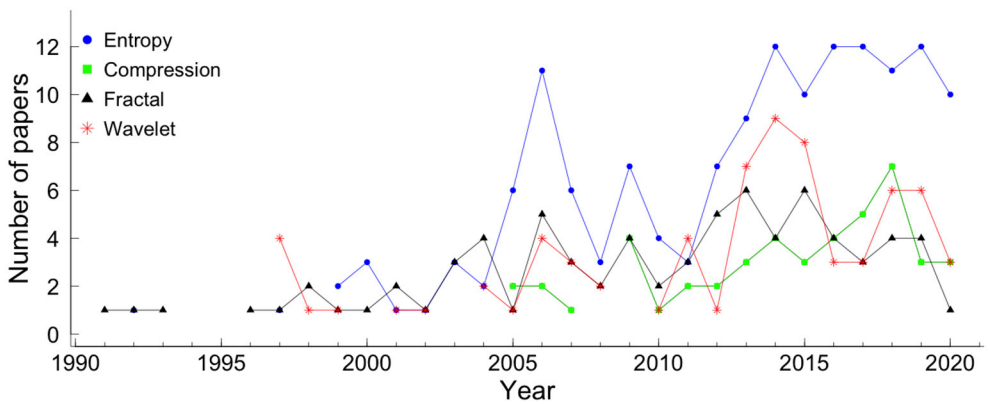

FIGURE 2 | The number of papers, covered by the systematic review, that applied each method. The blue circle represents entropy, the green square represents compression, the black triangle represents fractal analysis, and the red star represents wavelet.

42, 49, 50), Kolmogorov-Sinai entropy (30), cross-sample entropy (51, 52), tone-entropy (53), bubble entropy (47), and compression entropy (46).

3.1.1. Shannon Entropy (SE)

In 1948, Shannon (40) proposed the concept of entropy (Shannon entropy - SE) to measure how the information within a signal can be quantified with absolute precision as the amount of unexpected data contained in the message. The Shannon entropy is obtained by:

$$SE = -\sum_{i} p(x(i)) \cdot log\left(p(x(i))\right) \quad (1)$$

where $p(x(i))$ represents the probability of the point $x(i)$, of a time series $X = (x_1, x_2, ..., x_N)$.

Though SE was introduced back in 1948, and many new entropies appeared to overcome some of the SE limitations, some authors still applied it in the analysis of fHRV (46, 54).

3.1.2. Approximate Entropy (ApEn)

In 1991, Pincus et al. (55) developed a regularity statistic tool to quantify the amount of regularity and the unpredictability of fluctuations over time-series data. The ApEn is based on the assumption that healthy dynamic stability comes from specific networks' specific mechanisms and properties. When a vulnerable connection arises between systems or within one, it is

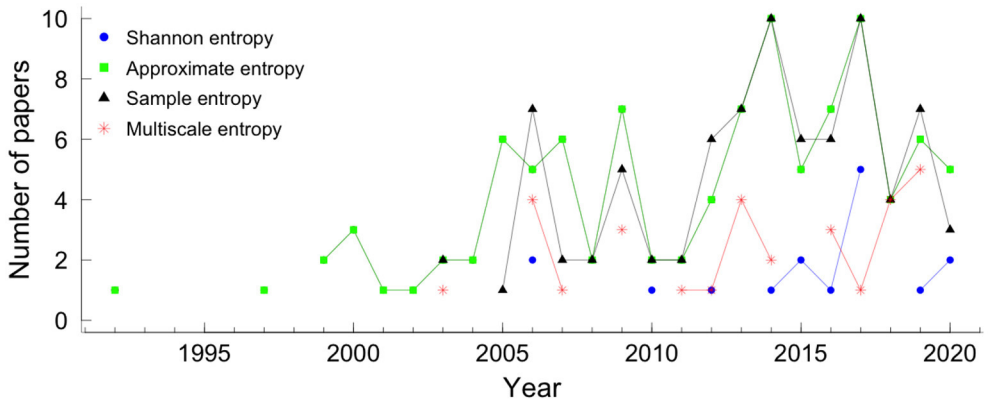

FIGURE 3 | Entropies by year. The colors and symbols represent the different measures of entropy. The blue circle, green square, black triangle, and star red represented Shannon entropy, approximate entropy, sample entropy, and multiscale entropy, respectively.

the disease mechanism, which is characterized by an increase of regularity of the series (56).

Considering a time series $X = (x_1, x_2, ..., x_N)$, in order to calculate the ApEn the new series of a vector of length m (embedding dimension), $X_m(i) = (x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+m-1})$ are constructed for each $i = 1, \ldots, N - m + 1$. For each vector $X_m(i)$, the value $C_m^r(i)$, where r is referred as a tolerance value, is computed as:

$$C_m^r(i) = \frac{\text{number of } d[X_i, X_j] \leq r}{N - m + 1}, \quad \forall j \tag{2}$$

Here, the distance between the vector $X_m(i)$ and its neighbor $X_m(j)$ is defined as:

$$d[X_m(i), X_m(j)] = max_{k=1,\ldots,m}|x(i + k - 1) - x(j + k - 1)| \tag{3}$$

Next, the average of the natural logarithm of $C_m^r(i)$ is computed for all i:

$$\Phi_m^r = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} ln(C_m^r(i)) \tag{4}$$

Since in practice N is a finite number, the statistical estimate is computed as:

$$ApEn(m, r) = \begin{cases} \Phi_m^r - \Phi_{m+1}^r & \text{for } m > 0 \\ -\Phi_1^r & \text{for } m = 0 \end{cases}$$

In the particular case of the ApEn, the most common value is $m = 2$. However, many algorithms were proposed to estimate the smallest sufficient embedding dimension, m. One of the most used methods is the "false nearest-neighbors" algorithm proposed by Kennel et al. (57). Though, the limitation of this method relies on the subjective definition of false neighbor (58). To overcome this limitation, Cao (58) proposed a new method.

For estimation of an appropriate time delay various approaches have been proposed. The most used two are the autocorrelation function and the average mutual information function (59). Pincus (60) and Pincus and Goldberger (61) recommends values between 10 and 25% of the standard deviation of the data, hence obtaining a scale-invariant measurement. The approach of choosing a fixed r value was also used with success (62, 63). However, the values of entropy in this case are usually highly correlated with the time series standard deviation. Lu et al. (64) showed that ApEn values varied significantly even within the defined range of r values and presented a new method for automatic selection of r that corresponds to the maximum ApEn value.

3.1.3. Sample Entropy (SampEn)

In 2000, Richman and Moorman (65) proposed the sample entropy (SampEn), with the same purpose as ApEn, to evaluate the randomness of biological time series, in particular, the HR time series. The main limitation of the ApEn is the dependence on the record length, i.e., the ApEn is lower for short records, and if one time series is higher than another, it should not remain higher for all conditions (65). In order to overcome the limitations, the authors proposed a new family of statistics, $SampEn(m, r)$, which, with some differences, reducing bias specially in short data sets:

1. self-matches are not counted;
2. only the first N-vectors of length are considered;
3. the conditional probabilities are not estimated in a template manner.

To calculate the value of SampEn (65) the parameters m, and r defined for ApEn are needed. Considering A as the number of vector pairs of length $m + 1$ having $d[X_m(i), X_m(j)] \leq r$, with $i \neq j$ and B as the total number of template matches of length m also with $i \neq j$, the SampEn is defined by the equation:

$$SampEn = -ln\frac{A}{B} \tag{5}$$

This probability measure is computed directly as the logarithm of conditional probability and not from the logarithmic sums ratio, showing relative consistency in cases where ApEn does not (65).

3.1.4. Multiscale Entropy (MSE)

ApEn and SampEn have the disadvantage of outputting a single index concerning the time series's general behavior, thus not revealing its underlying dynamics. MSE has been widely employed in the biomedical signal analysis as it allows measuring signal properties at different time scales (66, 67).

Considering a time series $X = \{x_i\}$ of N points, it constructs consecutive coarse-grained time series $y_j^{(\tau)}$, replacing τ non-overlapping points by their average. The MSE curve is created by computing the entropy for each of the scales and plotted vs. the scale. The information of the different time scales is clustered in the complexity index defined as the area under the MSE curve.

The estimation of the complexity methodology has to follow the multiscale application requirements, and SampEn was proposed using a tolerance r obtained from the original series and keeping it constant for all scales (67). Other authors were in favor of choosing an individual tolerance level r for each scale (68, 69). For example, the quadratic sample entropy permits a personalized estimation of r for each scale in short data (70).

The physiological interpretation of multiscale complexity is not always clear once, in a complex dynamic system, all scales might be affected by regulating influences (71). Low complexity scales indicate regular patterns with periodicity, but isolated ones would indicate one single frequency oscillation periodicity that usually is not present in complex systems. However, it is typical of the appearance of correlated neighboring scales (41, 67).

3.2. Compression

Dynamic systems theory was firstly linked with information theory by Kolmogorov (72), in 1958. Years later, "algorithm information theory" was then independently proposed by three different authors, Solomonoff (73), Kolmogorov (74) and Chaitin (75).

Let x be a finite length binary string, $\mathcal{U}$ be a universal computer, $l(x)$ denote the length of the string x and $\mathcal{U}(p)$ the output of the computer $\mathcal{U}$ when presented with a program p. The Kolmogorov (or algorithmic) complexity (KC) of a string x with respect to a universal computer $\mathcal{U}$, $K_{\mathcal{U}}(x)$, is defined as the shortest description length of x over all descriptions interpreted by computer $\mathcal{U}$. In different words, KC quantifies how "random" an individual object is in terms of the number of bits necessary to describe it. For a random string, the output of $K_{\mathcal{U}}(x)$ function will be the original string's length as any compression effort will end in information loss. The more reoccurring patterns, the less complex the signal is. Although this concept is objective, its applicability is limited to the fact that it is not computable. Compressors are a close upper-bounded approximation of the $K_{\mathcal{U}}(x)$ function. For over 30 years, data compression software has been developed for data storage and transmission efficiency purposes, and more recently, compression has been utilized in health research.

The innumerous compressors found in the literature can be divided into two big groups: lossless or lossy. The former group is composed of compressors in which, after being decompressed, all original information is restored. For the lossy group, this is not guaranteed, particularly for redundant information. The most applied compressors in health research belong to the first group.

The Lempel–Ziv algorithm was introduced, in 1976, by Lempel and Ziv (76) based on 'the concept of encoding future segments of the source output via maximum-length copying from a buffer containing the recent past output.' It was the starting point for different compressors such as the Lempel–Ziv–Markov chain algorithm, LZ77, LZ78, and gzip. The bzip2 was developed by Seward (77) and used the block sort algorithm giving speedy results.

In order to estimate the complexity of a physiological signal using compression, different approaches have been used, such as an increase/decrease coding system using a binary (30, 78, 79) or ternary alphabet (80, 81).

Compression also has been used for research purposes in a wide variety of fields such as literature (82), music (83), computer virus and internet (84) traffic, but only in 2004, it was first applied in HRV time series by Ferrario et al. (85). Here, compression demonstrated to differentiate healthy fetuses from unhealthy ones. In fact, the former group complexity calculated with LZ achieved similar results to random noise (meaning high complexity), while in the latter group, its complexity was lower, showing sinusoidal patterns. The applications of compression in health research range from event detection [such as epileptic seizure (86), the onset of ventricular tachycardia or fibrillation (87) and changes from sleep to waking state in-depth anesthesia (88)], characterizing neural spike trains (89), fHR biometric identification (90) or in DNA sequences studies (91). A distinct approach to applying compression on a time series uses the normalized compression distance (NCD) measure, a dissimilarity learning approach first used in fHR by Santos et al. (37).

From all 270 papers included in this review, 46 (17%) show results with compression. Its usage throughout recent years can be seen in **Figure 4**.

3.3. Fractal Analysis

Fractality indices quantify self-similarity and fractal- or multifractal-like behaviors. The heart rate fluctuates on different timescales and is similar to itself, which is a good premise for a fractal analysis approach (30).

Of all 270 papers included in this review, 28.1% applied fractal analysis. More specifically, 35 (13.0%) used detrended fluctuation analysis (DFA), 34 (12.6%) show results with fractal dimension (FD), 14 (5.2%) used Hurst exponent and 14 (5.2%) multifractal analysis (see **Figure 1**). **Figure 5** shows the number of papers that applied measures of fractal analysis by year.

Fractal dimension, Hurst exponent, and DFA are described in sections 3.3.1–3.3.3, respectively. The multifractal analysis describes more complex signals than those fully characterized by a monofractal model but requires many local and theoretically infinite exponents to characterize their scaling properties completely. The multifractal detrended fluctuation analysis (MF-DFA), the most applied multifractal method in the papers covered by the systematic review, is described in section 3.3.4.

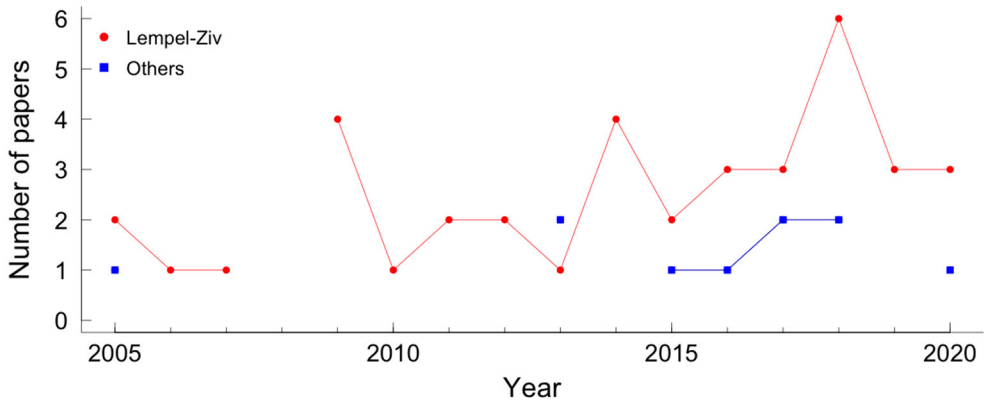

FIGURE 4 | Papers using compression in fetal heart rate, by year. The red circle represents the Lempel-Ziv compressor and the blue square represents the other compressors.

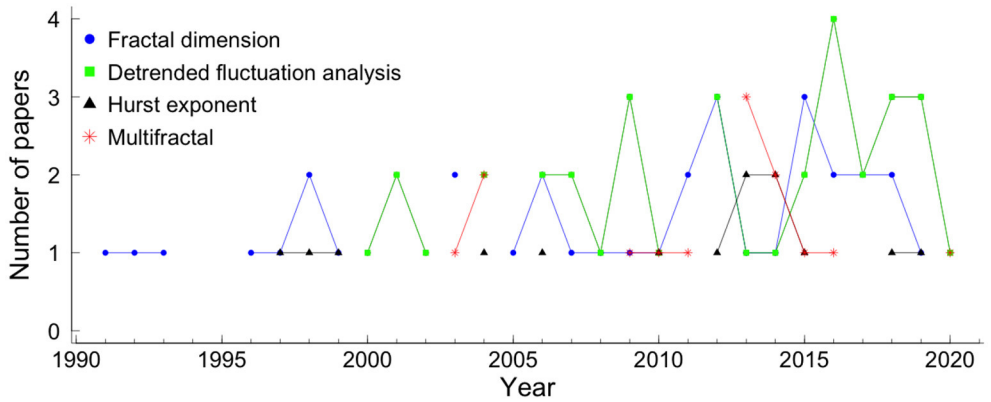

FIGURE 5 | Papers using fractal analysis in fetal heart rate, by year. The blue circle represents the fractal dimension measure, the green square represents the detrended fluctuation analysis, the black triangle represents the Hurst exponent, the red star represents the multifractal analysis.

3.3.1. Fractal Dimension (FD)

A fractal dimension (FD) is a statistical index of how the detail in a pattern changes with the scale at which it is measured. The FD appears to provide a measure of how much space an object occupies between Euclidean dimensions. The higher the FD value, the more irregular the signal is and, therefore, the more self-similar the signal will be.

Of the various algorithms available to calculate the FD of a time series, the four most used are the algorithms proposed by Katz (92) and Higuchi (93, 94), the correlation dimension, and the box-counting dimension (95). More details on the FD calculation algorithms of a time series can found at Henriques et al. (96).

3.3.2. Hurst Exponent

Hurst exponent (HE) or Hurst coefficient is a dimensionless estimator used to evaluate the self-similarity and the long-range correlation properties of time series (97). There are many algorithms to estimate the HE parameter in the literature. The oldest is the so-called rescaled range analysis (R/S) popularized by Mandelbrot and Wallis (98, 99) and it is defined in terms of the asymptotic behavior of the rescaled range (a statistical measure of the variability of a time series). Alternative methods to estimate HE include detrended fluctuation analysis (100, 101), periodogram regression (102), aggregated variances (103), local Whittle's estimator (104), first return method (105), wavelet analysis (106), both in the time domain and frequency domain. Furthermore, there is a relation between HE and the FD, given by $FD = E + 1 - HE$, where E is the Euclidean dimension, which for time series is 1 obtaining their relationship $FD = 2 - HE$ (107). The HE may range between 0 and 1 and can indicate:

- $0 < HE < 0.5$: time series has long-range anti-correlations;
- $HE = 0.5$: there is no correlation in the time series;
- $0.5 < HE < 1$: there are long-range correlations in the time series;
- $HE = 1$: the time series is defined self-similar, i.e., it has a perfect correlation between increments.

3.3.3. Detrended Fluctuation Analysis (DFA)

Detrended fluctuation analysis (DFA) quantifies intrinsic fractal-like (short and long-range) correlation properties of dynamic systems (101). Two advantages of DFA over conventional

methods (such as the HE method) are that this method allows the detection of intrinsic self-similarity embedded in a non-stationary time series and also avoids the detection of apparent self-similarity (108).

To execute the DFA algorithm the first step is to integrate the time series (of length N). The next step is to split the integrated time series into N_n windows of equal length n. Then, a least-squares line is fitted to the data, in each window of length n. The y-coordinate of the straight-line segments is given the name of $y_n(k)$. Then, the integrated time series is detrended, $y_n(k)$, in each window. The root mean square fluctuation of this integrated and detrended series is calculated by the following equation:

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2}. \quad (6)$$

This algorithm is repeated for all time scales (box sizes) to characterize the relationship between $F(n)$, the average fluctuation, and the box size, n. Normally, $F(n)$ increases with the size of the window, according to $F(n) \propto n^{\alpha}$. The α exponent can be viewed as an indicator of the "roughness" of the original time series: the higher the value of α, the smoother the time series:

- if $\alpha \simeq 0.5$, the time series represents uncorrelated randomness (white noise);
- if $\alpha \simeq 1$ (1/f-noise), the time series has long-range correlations and exhibits scale-invariant properties;
- if $\alpha \simeq 1.5$, the time series represents a random walk (Brownian motion).

Usually, the DFA method involves estimating a short-term fractal scaling exponent, α_1, and a long-term scaling exponent, α_2.

3.3.4. Multifractal Detrended Fluctuation Analysis (MF-DFA)

The multifractal DFA (MF-DFA) calculation (109, 110) is similar to the DFA since only two additional steps are taking into consideration. The fitting procedure in the MF-DFA can be linear, quadratic, cubic, or higher-order polynomials (MF-DFAm - the m^{th} order of the MF-DFA) (101, 111, 112). By comparing the results obtained for different MF-DFA orders, it is possible to estimate the order of the polynomial segment trends in the time series (109, 112). The procedure must be repeated for various n time scales, as we are interested in how this q-dependent fluctuation function depends on the n time scale for different q values. The other additional step is a q dependent averaging procedure obtaining a generalized (multifractal) scaling exponent $h(q)$. For $q = 2$, the standard DFA procedure is retrieved.

The main problem with the MF-DFA method is that all the steps are deeply dependent on the user's decisions. The Multiscale multifractal analysis (MMA) (71, 113) is a generalization of the MF-DFA method. The method creates a Hurst surface h(q,s), allowing a broader analysis of the fluctuation properties and more stable results. Also, all multifractal methods, including MMA, require a relatively long time series to analyze.

3.4. Wavelets Analysis

The first appearance of the term wavelet was in an annex to Haar thesis' (114). However, it is considered that the wavelet theory was developed in the late 1980s by Mallat (115), Daubechies and Bates (116, 117) to meet the needs for adaptive time-frequency analysis applied to signal processing, mathematics, physics, and engineering. Wavelets are functions that satisfy a series of mathematical parameters and are used in the representation of data or other functions. The term wavelet comes from the fluctuation around the axis, integrating to zero (the areas above the axis and below are the same). Wavelet algorithms process information at different scales (or resolutions). The decomposition of a function using wavelets is known as a transformed wavelet, and it has continuous and discrete variations. Due to the ability to decompose functions in frequency and time domains, wavelet functions are powerful tools for signal processing, widely used in data compression, noise elimination, separation of components in the signal, identification of singularities, and auto-similarity detection.

Let $\psi_{s,u}(t)$, s, $u \in \Re$, $s > 0$ be a family of functions defined as translations and re-scales of a single function $\psi(t) \in L^2(\Re)$, $L^2(\Re)$ denotes the space of square-integrable functions on $\Re$ (118),

$$\psi_{s,u}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \quad (7)$$

where s is the scaling parameter and u the position parameter. The parameter u indicates that the function $\psi(t)$ was translated on the t axis (translation parameter) by a distance equivalent to u. The parameter s causes a scale change, increasing (if $s > 1$) or decreasing (if $s < 1$) the wavelet formed by the function. The wavelet is defined as a mother wavelet $\psi(t)$ [equivalent to $\psi_{1,0}(t)$], with a family of scale and time daughter wavelets $\psi\left(\frac{t-u}{s}\right)$. Therefore, daughter wavelets constitute a family of curves with a shape identical to that of the mother wavelet, displaced in time and scaled in amplitude. In the time domain, the wavelet transform measures the correlation between the $f(t)$ signal and the daughter wavelets.

The wavelet $\psi_{s,u}(t)$ has the following basic properties:

$$\int_{-\infty}^{\infty}\psi(t)dt = 0 \text{ and } \int_{-\infty}^{\infty}|\psi(t)|^2dt = 1. \quad (8)$$

The wavelet $\psi_{s,u}(t)$ has to meet the admissibility condition for the transformation to be invertible (116).

The term $\frac{1}{\sqrt{s}}$ is a normalization factor that ensures that the energy of $\psi_{s,u}(t)$ is independent of s and u, such that:

$$\int_{-\infty}^{\infty}|\psi_{s,u}(t)|^2dt = \int_{-\infty}^{\infty}|\psi(t)|^2dt \quad (9)$$

The continuous wavelet transform (CWT) of signal $f(t)$ is defined as:

$$W_{\psi}f(s,u) = \langle f(t), \psi_{s,u}(t)\rangle = \frac{1}{\sqrt{s}}\int_{-\infty}^{\infty}f(t)\psi\left(\frac{t-u}{s}\right)dt. \quad (10)$$

The CWT coefficients $W_\psi f(s,u)$ can be obtained by continuously varying the scale parameter s and the position parameter u. For real discrete signals $f(n)$, as is the case for the fHR signal, $W_\psi f(s,u)$ can be calculated according to

$$W_\psi f(s,u) = \frac{1}{\sqrt{s}} \sum_{n=1}^{N} f(n) \psi \left(\frac{t-u}{s} \right). \quad (11)$$

If s is a continuous variable then $W_\psi f(s,u)$ is called the continuous wavelet transform. However, if $s = a^j$ and $u = n * u_0 * a^j$ where the integers j and n control the wavelet dilation and translation respectively; a is a specified fixed dilation step parameter set at a value greater than 1, and u_0 is the location parameter which must be greater than zero then $W_\psi f(s,u) = W_\psi f(j,u)$ is called the discrete wavelet transform (119). A useful property of the wavelet transform is that it can be viewed as the application of a filter bank (each filter corresponds to one scale) (120). Some authors, such as, Zhao et al. (121) and Papadimitriou et al. (122) apply different scale values, but, in practice, $s = 2^j$ and $u_0 = 1$ are the most popular scale in fHR analysis (123–125).

There are a vast number of different mother wavelets, each suitable for different applications. In particular, several wavelet families have been proposed for fHR analysis. From all 270 papers included in this review, 26.3% applied wavelet analysis. The Daubechies (36 papers), spline (11 papers), symlets (11 papers), and coiflet (8 papers) wavelet families were the most applied in fHR analysis (see **Figure 1**). The application of wavelet analysis in fHR has intensified in the last 10 years, **Figure 6**.

4. RESULTS

There is no doubt of the importance of non-linear measures in fetal monitoring, as they enrich the signal description by providing new indicators for classification and diagnostic purposes. Numerous studies have documented the changes in fHRV during gestation, and fetal growth is associated with a drop in fetal heart rate and increased variability. As non-linear measures started being used, authors started to link their association with different physiological regulatory systems.

The history of non-linear methods reported to fHR summarizes 30 years. However, in the last 15 years, there has been a notable increase in their application to study fHR dynamics (see **Figure 7**). The main research objectives covered by this systematic review were signal processing (60 papers), hypoxia (56 papers), maturation or gestational age (53 papers), IUGR (44 papers), and fetal well-being or fetal distress (26) (see **Figure 7** and **Table 2**). Also, in **Figure 7**, the evolution of papers' of the five most cited research objectives is presented per year.

Hypoxia can be caused by prolonged or profound asphyxia, an oxygen deficiency due to a pathological change in either fetal or maternal components of the placenta, when there is an exchange of carbon dioxide and oxygen by the fetus during labor. This state leads to an accumulation of carbon dioxide leading to fetal acidemia, resulting in a lower pH in the fetal blood vessels. Early detection of which babies are at risk of acidemia is crucial, as it decreases the chance of a post-diagnosis of cerebral palsy, neonatal encephalopathy, or even death (126). To relate fHR with umbilical artery pH is, therefore, of extreme importance. However, the proper definition of fetal acidemia is still not established as different authors consider different pH cutoffs. Moreover, some authors also include in the definition the value of the base excess or base deficit (127). Some authors defined as "at risk of acidemia" when $pH < 7.20$ (33, 47, 128–136) or $pH < 7.15$ (30, 121, 137–143); others define when $pH < 7.1$ (43, 126, 144–146) or even when $pH < 7.05$ (38, 44, 48, 78, 147–165). Some studies used clinical experts to identify episodes of hypoxia and asphyxia, such as, (54, 166, 167). Another challenge relate to this pathology is to collect enough data for a proper acidemia analysis since prevalence of an acidemic fetus ranges from 0.6 to 3.5% (168, 169). From the 56 papers that aim to analyze of perinatal hypoxia or asphyxia 40 papers applied entropies, 19 applied compression measures, 23 applied fractal measures, and 23 papers applied wavelets analysis.

The development of non-invasive ultrasound techniques allowed a better estimation of gestational age and, therefore, the definition of a crucial fetal outcome: small for gestational age (SGA), which corresponds to fetuses having a weight lower than the 10th percentile adjusted to gestational age. Nevertheless, healthy babies can also be considered SGA, so it is still a challenge to decide whether the small dimensions are due to physiological or pathological conditions (170). Related to SGA fetus, one of the most common pathologies is IUGR. IUGR is a metabolic dysfunction inhibiting the fetus from achieving its average size. With a prevalence of $5 - 8\%$ in the general population, it can complicate 10–15% of all pregnancies (171). IUGR is the second cause of perinatal mortality, after prematurity (172), and is still an important challenge for diagnosis and management (173). From the 44 papers that aim to study IUGR, 32 papers applied entropies, 15 applied compression measures, 11 fractal and only 2 papers used wavelets.

The effect of an antepartum vs. intrapartum analysis on the complexity indices and the differences in the signal acquisition methods are important to correctly evaluate and assess fetus well-being (174, 175). Throughout pregnancy, the fetus interacts with its environment, as the mother sets the framework for the state and development of the fetus (176). In a study where the mother's breathing was controlled, Van Leeuwen et al. (176) found that the presence or absence of interaction between mother and fetus cardiac activity might be due to maternal respiration. Also, the fetal cardiac system seems to have the capability to adjust its activation rate when responding to external stimuli. Spyridou et al. (177) studied the effect of smoking in fHR and found differences with several linear and non-linear parameters (such as, mutual information, MSE, and compression). In particular, it was shown less complexity for fetus exposed, enhancing its danger. From the 26 papers that aim to study fetal well-being or fetal distress 17 papers applied entropies, 4 applied compression measures, 7 fractal, and only 3 papers used wavelets.

When assessing fHRV, it is essential to control any factor which might confound its interpretation. Some of the most studied factors are the baby's maturation reflected in gestational age, behavioral state, and maternal condition (178). From the

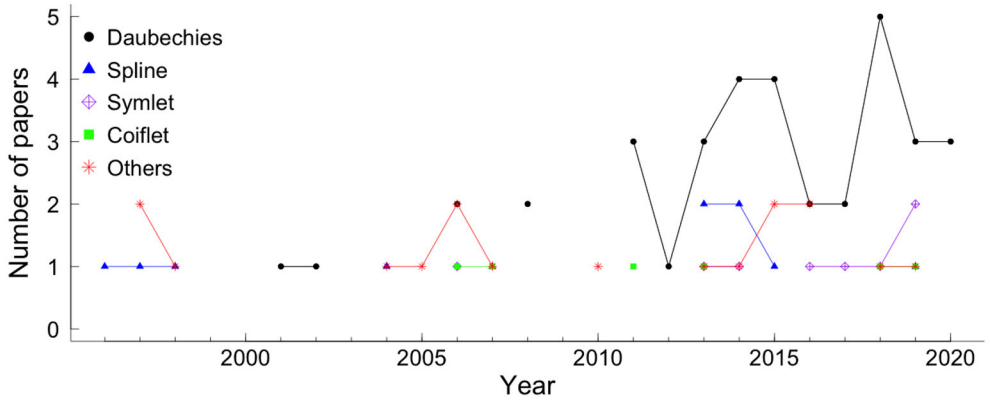

FIGURE 6 | The wavelet families most used in fHR analysis, by year. The colors and symbols represent the wavelet families. The black circle, blue triangle, purple diamond plus, green square, and star red represented Daubechies family, spline family, symlet family, coiflet family, and other families, respectively.

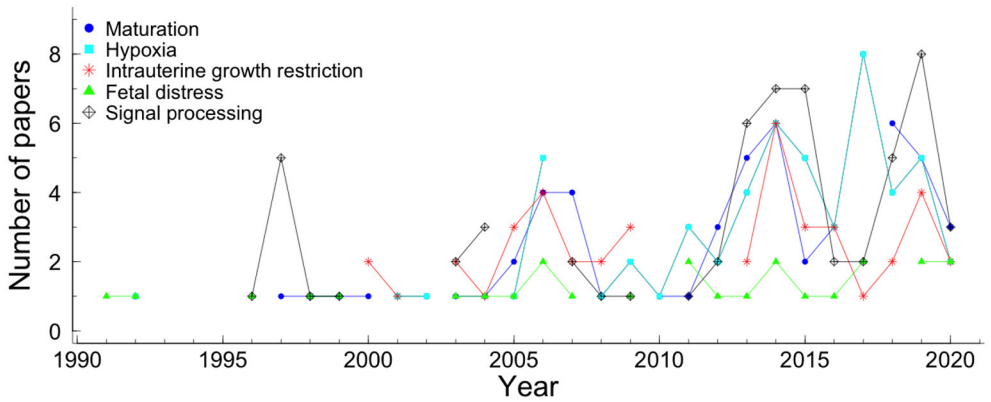

FIGURE 7 | The five most cited research objectives of the papers that applied non-linear methods in fHR, by year of publishing. The colors and symbols represent the different research objectives. The blue circle, cyan square, star red, green triangle, and black diamond plus represent maturation, hypoxia, intrauterine growth restriction, fetal distress, and signal processing, respectively.

53 papers that aim to study maturation or gestational age 38 papers applied entropies, 7 applied compression measures, 18 fractal and only 3 papers used wavelets. First trimester observations during pregnancy have shown a low intraindividual variation of the fHR, compared to variation between different fetuses (179). Later in pregnancy, Arduini also found this high intraindividual consistency concerning fetal behavioral states' characteristics, particularly fHR, in 2 consecutive days (180). In fact, an association between individual differences in prenatal heart rate and HRV and postnatal neural development has been reported (181). Besides these factors, Gonçalves et al. (131) and Spyridou et al. (182) noted that gender also has an effect on fHR analysis and should be considered, while Tagliaferri et al. (183) found differences on both linear and non-linear indices between different ethnic groups. Gender was also shown to influence maternal heart rate (MHR) (52). Even when twins are considered, sex differences were found both by linear and non-linear indices (184). Fetal presentation at birth has also been studied (185, 186). Reports are stating that breech fetuses have worse neurological outcomes compared to cephalic presentation ones (187, 188). Furthermore, in a study by Choi and Hoh (189), non-linear dynamic indices were able to differentiate normal pregnancies from ones with partial placental abruption with high accuracy, while linear indices were not.

The evaluation of neonatal behavior has shown more success in predicting neurodevelopment disability than neurological examination (190). Therefore, the same approach was adopted for fetal well-being assessment. These fetal behavioral states were introduced back in 1982 in studies combining the assessment of fetal body and eye movements (191). They include calm or non-eye movement sleep state (1F), active or rapid eye movement sleep state (2F), calm wakefulness state (3F), and active wakefulness state (4F). The importance of these definitions in understanding fetal physiology, interpretation of fHR monitoring, and diagnosis of pathological conditions is described with more detail elsewhere (192). There are associations between fetal behavioral states and fHR patterns. 1F is related to a stable baseline

TABLE 2 | Research objectives and non-linear methods of the papers selected.

	Non-linear analysis													
	Entropy					Comp		Fractal				Wavelet	O	T
Objectives	**ApEn**	**SampEn**	**SE**	**MSE**	**O**	**LZ**	**O**	**FD**	**HE**	**DFA**	**O**			
Healthy														
Maturation	19	12	4	13	11	7	1	3		12	4	3	14	53
Activity/behavior	10	5	1	2	3	1				3	1	1	3	14
Gender	10	8		2	2	1		1		1			2	11
Presentation	3	2				1		1						3
RM										1				1
Labor	3	2	1	1		1	1			1		1	3	6
Cesarean		1			1			1		1		1	3	4
Preterm	2	2		1	1	1		1		1		1		3
Twins	1	1						1					1	2
Nuchal cord	1													1
FCTE													1	1
Self-organization			1										1	1
Ethnic origins	1													1
Pathologies														
Hypoxia	29	28	7	6	11	14	5	11	6	6	7	23	27	56
IUGR	26	13	2	6	3	15		2	1	7	1	2	16	44
Fetal distress	13	7	2	3	4	3	1	4	3	1	2	3	12	26
SIDS	1	1												1
Intrauterine demise				1										1
PPA	1	1						1		1				1
Anencephalus										1			1	1
Maternal														
MP	5	1		1		1		1		2	1	2		9
Hypnosis	1	1	1											1
Steroid treatment					1									1
Uterine contraction	2	2	1	1									4	4
Signals														
MFCC	1		3		4	1	1						4	7
BI					1		1						1	1
FCEC					1								1	1
Signal Processing	3	7					1	4	2	2		45	28	60
Signal acquisition	1	1												1
fHR baseline										1			4	4
Others														
Expert annotation	3	4		1	1	3		4	2	3			4	6
Patterns	5	3						1	1	1			3	9
Fractal value					1			1						1
Total	101	82	16	30	38	38	8	34	14	15	15	71	112	

ApEn, approximate entropy; BI, biometric identification; CD, correlation dimension; Comp, Compressor; DFA, detrended fluctuation analysis; FCEC, fetal cardio-electrohysterographic coupling; FCTE, fetal cardiac timing events; FD, fractal dimension; HE, Hurst exponent; IUGR, intrauterine growth restriction; LZ, Lempel-ziv; MFCC, maternal-fetal cardiac coupling; MSE, multiscale entropy; MP, maternal pathologies; O, others; RM, Respiratory movement; PPA, partial placental abruption; SampEn, sample entropy; SE, Shannon entropy; SIDS, sudden infant death syndrome; T, total.

with absent or sporadic and short-lasting accelerations; 2F is associated with a stable baseline and frequent accelerations, and it is the most frequent state. 3F is rare and is usually very short in time. It also has a stable baseline but with wide variability and no accelerations. 4F shows repetitive and long-lasting accelerations with eventual returns to the baseline (193).

This field's interest is not only focused on fHR tracings classification. Features like frequency and amplitude traditionally characterize physiological signals. However, these parameters do

not provide us with an insight into the regulatory processes underlying the signal dynamics, thus requiring a further extraction of more appropriate features, which has become a difficult task. These difficulties lie in the lack of a priori information on which process belongs to each component (i.e., fetal, maternal, or environmental) and the lack of knowledge on how each component behave (194). Much effort has been put into the signal acquisition and processing models because the extracted features' usability highly relies on the preprocessing steps' quality, such as artifacts removal, interpolation method, segmentation, and detrending signal (30).

In 2013, an open challenge was created, the Physionet Challenge (195), in order to promote the development of advanced signal processing techniques. Many different approaches were suggested, such as wavelet de-denoising, subspace decomposition and reconstruction, adaptive filtering and averaging, matched filtering, and entropy. Most of them followed these four steps: signal processing, maternal heartbeat detection, maternal heartbeat cancelation, fetal heartbeat detection. More information can be found in Di Maria et al. (196). These non-linear methodologies have been studied and applied to retrieve signal with the best quality possible, dismissing as much noise as possible. The preprocessing is even more important when adopting low-cost systems for signal extraction, as is the case of the fetal phonocardiography, which has a poor signal-to-noise ratio (11). From the 60 papers that aim to study signal processing 7 papers applied entropies, 1 applied compression measures, 7 fractal, and 45 papers used wavelets.

Table 2 presents the number of articles that applied each non-linear method for each research objective. Entropy, compression, and fractal measures are most used in classification papers, mainly when applied to analyze the variability of fHR in hypoxia, IUGR, and fetal distress. However, these measures are still underused in studies whose research objective is signal processing. On the other hand, wavelet analysis is most used when the research objective is signal processing (43 papers) or hypoxia (25 papers).

5. DISCUSSION

The number of articles probing the use of non-linear measures to assess the fHR signals analysis has been growing in the past decade. Non-linear analysis has been successfully applied in the study of fetal heart rate with several research objectives, such as fetal maturation or gestational age (197–199), fetal gender (182, 200), labor stages (201, 202), cesarean section (51, 203), preterm birth (80, 204), impact of nuchal cord on antenatal (205), fHR baseline (206), behavioral state (207, 208), IUGR (209–213), hypoxia (128, 137, 214, 215), fetal distress (216–218), maternal pathologies (219, 220), and signal processing (198, 221–224). Therefore, it is important that the scientific community is aware of the non-linear methods used depending on the research objective. Additionally, they are not yet used in clinical practice due to some critical concerns that need to be further discussed.

Systems, such as, Omniview SisPorto (225), OxSys (226), NST-Expert, which later became CAFE (227) already automatically deal with CTG assessment. All the fHR processing and analysis in these systems are based on morphological features defined by FIGO guidelines. In some, the CTG is complemented with the ST-analysis method. It has been shown that it slightly improves labor outcomes, but its use is not always possible since it requires an invasive measurement (228). However, none of these systems still integrates non-linear indices, so they can and should be optimized.

When analyzing fHR time series automated, there are two main aspects to contemplate: the signal properties and quality and the clinical characteristics that might influence the measures used. Accordingly, we found that the most studied research objectives in fHR are signal processing, hypoxia, and maturation. Furthermore, the results show that entropy is the most applied measure in fetal heart rate, followed by fractal analysis, wavelet analysis, and the least applied is compression. Although the application of entropy methods stands out, we can see that compression and wavelet analysis methods have been increasingly used in recent years. Also, highlighting the fact that entropy is the oldest method, and that is it has been extensively studied and refined when applied to much different time series (229). On the other hand, wavelets are widely used in signal processing (124, 222) dealing with the signal itself, handling problems such as noise (230) and frequency.

Routinely, the fetal heart rate monitors acquire the beat-to-beat intervals in milliseconds either from Doppler or electrocardiographic signals and then convert them to provide a sequence of instantaneous heart rates in beats per minute (bpm). However, when data is exported, it is sampled, implying an interpolation of signals (132). The sampling rate does not seem to affect many linear parameters, but differences were found when non-linear indices were considered (175). Caution must be taken when defining reference values for irregularity indices, such as entropy, as they depend on the sampling frequency, as shown in (175), where 2 vs. 4 Hz sampling was compared. It is most important not to compare computerized systems for heart rate frequency analysis that use different sampling rates (225, 231).

Several linear methods have been studied as a forecaster of fetal well-being by measuring the interaction between the fetal sympathetic and parasympathetic nervous systems and its effects on fetal cardiovascular activity (232). As the parasympathetic nervous system is more responsible for variations in short-term variability (STV), which usually assesses the beat-to-beat differences, it might be reduced in central nervous system hypoxia/ acidosis. If hypoxia is sustained and increases in severity, it leads to the loss of long-term variability (LTV) (233), resulting in a global decrease of sympathetic and parasympathetic activity. On the other hand, it has been shown that fetal hypoxia's early effects increased short and long-term variability (234). Notwithstanding, many studies verify the weakness of STV and LTV indices in identifying fetal pathologies (235). Furthermore, with fetus maturity throughout pregnancy, an increase in fetal autonomic nervous system activity and the sympathovagal balance is expected. Moreover, motor and neurological delay, as well as damage in specific brain areas with cognitive effects, also affect the STV (236). The IUGR showed a reduction in both components of the autonomic nervous system activity, which

modulates heartbeat intervals receiving inputs from the heart, the lungs, and the blood vessels (204, 209, 237).

The indices presented in this review are closely related to fetal heart variability. For instance, in (199) the authors showed that the complexity indices correlate highly with abnormal STV. In (143) the authors also report correlations between the complexity indices ranging from 0.53 to 0.78. Therefore, many studies found a reduction of complexity in the fHR signal associated with hypoxia/ acidosis. However, these indices were not always able to identify fHR from IUGR fetuses (237). Contrastingly, the fractal indices are measures of long-range correlations and long-term memory of time series, therefore, applied mainly in maturation studies.

Furthermore, many of the fetal heart rate analysis methods rely upon stationarity properties like mean, variance, and correlation structure. However, it is known that these fHR properties vary in time through events like uterine contractions. One way to counter this is to select small temporal windows where this property holds. Usually, an interval of 10–20 min is considered the minimum time window to perform the analysis for tracing classification and clinical decision (231, 238). In addition, many of the described measures are parametric measures. The choice of the ideal parameter is far from established in most cases. This heterogeneity limits the possible comparison between the results of different studies. In fact, in various papers, the choice of parameters is neither discussed nor even fully described.

Factors like fetus maturation, behavioral states, and maternal conditions are critical for a good assessment of the fetus and fully understanding their influence in the fHRV is no easy task. Incorporating such variables in predictive models for fetal evaluation will elucidate the importance of individual fHRV and increase its accuracy (178). Maternal psychological conditions such as stress and anxiety influence fHR and maternal hormones transferred via placenta or changes in the oxygen and nutrition supply for the fetus (239, 240). As seen in some results, gender is also a factor that should be taken into consideration. Even when twins are considered, sex differences were found both by linear and non-linear indices (184). Although Park et al. (185) found no significant differences between fetuses with different fetal presentation using spectral and complexity measures such as Lempel-Ziv complexity, ApEn, SampEn, and CD. Gonçalves et al. (186) found differences not only using linear indices but also with non-linear and spectral ones. This example of contrasting results reflects the difficulty and complexity of the fetal assessment. In this case, and according to the authors, the discrepancy might have resulted from different inclusion criteria, conditions for fHR recording, the occurrence of maternal fasting, time interval between acquisition and delivery, and equipment used. Moreover, an interesting study comparing uterine contraction influence on fHRV features between acidemic and non-acidemic fetuses suggested that separating contractions from rest periods improves fHRV analysis in detecting asphyxia during labor (151).

Having as a premise that humans are a result of self-organization and adaptation process and that ontogenetic development reflects phylogenetic development and indices of developmental biology may be helpful in fetal maturation assessment. Many studies addressed here found HRV changes, such as variability increase and pattern formation (204). These universal developmental features deliver appropriate measures of fetal maturation. Therefore, it seems only natural that these self-organization and adaptation features might better understand and identify developmental disorders (241). In fact, attention-deficit hyperactivity disorders in teenage boys were associated with antenatal maternal anxiety (242), which might influence fetal humoral development and autonomic control reproduced in heart rate patterns. This phenomenon, resulting from adverse influences on the fetus explained by epigenetic mechanisms, is called "fetal programming." Therefore, early identification of fetal developmental disorders is essential as they may not be wholly compensated for later postnatal therapies (243). Many different approaches to fHR processing and analysis have been studied. They range from simple feature extraction methods to more sophisticated classification programs and joining research centers from different countries for joint projects, as the Digi-Newb project (244). Usage of continuous non-invasive evaluation, such as the usage of wearables, have been discussed (27, 245) and will contribute to the patient's care improvement since it will improve data gathering, reducing costs of fetal monitoring. Insurgent approaches are opening new windows on the continuous monitoring of fetal development. A single index cannot retrieve all the information from pathophysiological processes in the fetus's development, so approaches considering both linear and non-linear measures, through multivariate analysis, can improve the assessment of both fetal and maternal well-being.

In (35, 246), time, spectral and complexity indices were used as parameters to discriminate fetuses who were or not in a distressed state. Ferrario et al. (247) conclude that compression quantifies the rate of new patterns arising as the signal evolves, whereas entropy quantifies the recurrence of repetitive patterns. This idea of complementary of different indices is also supported in other papers (130, 145, 167, 247). It seems only logical for such a complex/chaotic system to be evaluated using a multiparameteric approach through advanced classification techniques capable of discriminating fetuses in distress in non-linear regions of a multidimensional space (30). With this approach, Signorini (248) was able to classify IUGR fetuses with accuracy, sensitivity, and specificity above 90%.

Mapping from feature space captured from the fHR signal to the space of decision or diagnosis, many machine learning, and deep learning techniques has been applied. Some examples are: support vector machines (38, 150, 152, 249, 250), conventional methods like k-nearest neighbors (250, 251), a hybrid approach using grammatical evolution (146, 252), artificial neural networks (134), and random forests (49, 253, 254). Cömert and Kocamaz (166) introduced a novel software for comprehensive CTG signals analysis, named CTG Open Access Software (CTG-OAS). This software embeds machine learning tools, such as preprocessing, feature extraction, feature selection, and classification. Fergus et al. (51) demonstrated, using deep learning tools, that machine learning significantly improves the efficiency of detecting cesarean section and vaginal

deliveries, compared with the usual visual assessment. In this paper, impressive results were achieved, with both sensitivity and specificity over 90%. One problem of comparing these classification approaches is the apriori definition of the classes. For example, as said before, the definition of acidemia based on the umbilical cord artery's pH varies greatly between studies. Karvelis et al. (255) proposed a classification approach based on weighted voting of clinical annotations. These weights are estimated by using a latent class model with three or four latent classes. Moreover, as these learning techniques depend on the signal and the linear and non-linear measures computed, all the previously referred concerns must be contemplated meticulously. Therefore, the machine and deep learning techniques are particularly resourceful when the measures are thoroughly probed and understood. In this systematic review were found several other articles with machine learning and deep learning techniques. However, the description of these techniques is not the focus of this paper. Future work that analyzes in detail the machine learning and deep learning techniques that apply measures based on fHR dynamics should be considered.

The non-linear methods described in this review are entropy (Shannon, approximate, sample, and multiscale), compression, fractal analysis (fractal dimension, Hurst exponent, detrended fluctuation analysis, and multifractal detrended fluctuation analysis), and wavelet analysis. Other non-linear methods were found in our review, such as Poincaré plot (217), symbolic dynamics (256), phase rectified signal average (210, 211, 257, 258), Lyapunov exponents (259), and recurrence plot analysis (137). In the recent years, Phase Rectified Signal Averaging and derived parameters have been largely applied in fHR analysis to face the problem of accelerations and deceleration which are characteristic of the fHR signal.

Due to the high heterogeneity of study designs, data acquisition methods, aims of the studies, signal processing techniques, and measures (and parameters) used, no meta-analysis was possible to be performed.

This systematic review confirmed the importance of non-linear fetal monitoring measures to analyze the fetus' well-being and pathologies' prediction. The methods probed successfully diagnose pathologies, and new techniques are being proposed and explored to improve that prediction. However, the contradictory results of some of the findings due to the characteristic of the signal, or the sensibility of the measures to some clinical factors, such as fetus sex and gestational age, revealed that the use of these findings in clinical practice is far from reality. These results inhibit the reach for a gold standard or the creation of a decision support system. This review determined the significance of creating several small meta-analyses that might focus on a specific research aim. Additionally, a sizeable multicentric study that can assess the multitude of perspectives involved in the fHR signal analysis is imperative.

6. CONCLUSIONS

Non-linear measures based on the concepts of chaos, fractality, and complexity gained space in the analysis of fetal heart rate. Good results were achieved in signal processing, in the analysis of fetal well-being, and in diagnosing and predicting pathologies. This systematic review of the non-linear methods (entropy, data compression, fractal analysis, and wavelet analysis) applied to fetal heart rate dynamics includes 270 papers. The application of non-linear methods in the fHR analysis is around 30 years old. However, its application has significantly increased in the last 15 years. This review's main contributions are a detailed description of the non-linear methods most applied in the fHR papers and a discussion of the research objectives. Signal processing, hypoxia, and maturation lead the research objectives of papers that use non-linear analysis in fHR. We found that entropy has been the most used method in classification analysis. Despite, in signal processing, the most used method is wavelet analysis. Machine learning and deep learning techniques should also be analyzed with results in the study of fHR dynamics using linear and non-linear measures. The multitude of conditioning involved in the analysis and classification of the fHR, from the signal characteristics to the effect of some clinical factors in the measures, limits the use of the non-linear measures in clinical practice and difficult the creation of a decision support system. Future studies should focus on a research question and perform a meta-analysis, probing the indices' performance.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JM-S, LA, CC-S, and AT conducted the preliminary literature review by establishing the search method and keywords. MR, JM-S, and AT determined the eligibility, summarized the findings from each study, and compiled them in tables. LC, CC-S, and TH analyzed the disagreements between reviewers. MR, JM-S, LC, AT, and TH writing and editing of the manuscript. CC-S, LA, AT, and TH supervision. All authors revised the paper critically for important intellectual content, made substantial contributions to the conception and design of the article, and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge professor Luís Azevedo for helping to build the queries for the systematic review. MR acknowledges Fundação para a Ciência e a Tecnologia (FCT) under scholarship SFRH/BD/138302/2018. The authors also acknowledge FCT, within CINTESIS, R&D Unit (reference UIDB/4255/2020).

REFERENCES

1. Cousens S, Blencowe H, Stanton C, Chou D, Ahmed S, Steinhardt L, et al. National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. *Lancet*. (2011) 377:1319–30. doi: 10.1016/S0140-6736(10)62310-0
2. Ayres-de Campos D, Bernardes J, Costa-Pereira A, Pereira-Leite L. Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. *BJOG Int J Obstetr Gynaecol*. (1999) 106:1307–10. doi: 10.1111/j.1471-0528.1999.tb08187.x
3. Bernardes J, Ayres-de Campos D. The persistent challenge of foetal heart rate monitoring. *Curr Opin Obstetr Gynecol*. (2010) 22:104–9. doi: 10.1097/GCO.0b013e328337233c
4. Spencer J. Role of cardiotocography. *Brit J Hospital Med*. (1992) 48:115–8.
5. Martinek R, Kahankova R, Jezewski J, Jaros R, Mohylova J, Fajkus M, et al. Comparative effectiveness of ICA and PCA in extraction of fetal ECG from abdominal signals: toward non-invasive fetal monitoring. *Front Physiol*. (2018) 9:648. doi: 10.3389/fphys.2018.00648
6. Sameni R, Clifford GD. A review of fetal ECG signal processing; issues and promising directions. *Open Pac Electrophysiol Therapy J*. (2010) 3:4. doi: 10.2174/1876536X01003010004
7. Bartlett M, Murray A, Dunlop W. Is fetal heart rate monitoring sufficiently sensitive to detect changes during labour? *J Biomed Eng*. (1992) 14:431–4. doi: 10.1016/0141-5425(92)90090-8
8. Parer JT. *Handbook of Fetal Heart Rate Monitoring*. Philadelphia, PA: Saunders (1997).
9. Kovács F, Horváth C, Balogh ÁT, Hosszú G. Fetal phonocardiography-past and future possibilities. *Comput Methods Prog Biomed*. (2011) 104:19–25. doi: 10.1016/j.cmpb.2010.10.006
10. Adithya PC, Sankar R, Moreno WA, Hart S. Trends in fetal monitoring through phonocardiography: challenges and future directions. *Biomed Signal Process Control*. (2017) 33:289–305. doi: 10.1016/j.bspc.2016.11.007
11. Chourasia VS, Tiwari AK. Fetal heart rate variability analysis from phonocardiographic recordings. *J Mech Med Biol*. (2011) 11:1315–31. doi: 10.1142/S0219519411004174
12. Persico N, Moratalla J, Lombardi C, Zidere V, Allan L, Nicolaides K. Fetal echocardiography at 11-13 weeks by transabdominal high-frequency ultrasound. *Ultrasound Obstetr Gynecol*. (2011) 37:296–301. doi: 10.1002/uog.8934
13. Zhang D, Zhang Y, Ren W, Sun F, Guo Y, Sun W, et al. Prenatal diagnosis of fetal interrupted aortic arch type A by two-dimensional echocardiography and four-dimensional echocardiography with B-flow imaging and spatiotemporal image correlation. *Echocardiography*. (2016) 33:90–8. doi: 10.1111/echo.12996
14. Quartero H, Stinstra J, Golbach E, Meijboom E, Peters M. Clinical implications of fetal magnetocardiography. *Ultrasound Obstetr Gynecol*. (2002) 20:142–53. doi: 10.1046/j.1469-0705.2002.00754.x
15. Grimm B, Haueisen J, Huotilainen M, Lange S, Leeuwen PV, Menendez T, et al. Recommended standards for fetal magnetocardiography. *Pac Clin Electrophysiol*. (2003) 26:2121–6. doi: 10.1046/j.1460-9592.2003.00330.x
16. Jaros R, Martinek R, Kahankova R. Non-adaptive methods for fetal ECG signal processing: a review and appraisal. *Sensors*. (2018) 18:3648. doi: 10.3390/s18113648
17. Hoyer D, Zebrowski J, Cysarz D, Gonçalves H, Pytlik A, Amorim-Costa C, et al. Monitoring fetal maturation-objectives, techniques and indices of autonomic function. *Physiol Measure*. (2017) 38:R61. doi: 10.1088/1361-6579/aa5fca
18. Steer PJ. Has electronic fetal heart rate monitoring made a difference? In: *Seminars in Fetal and Neonatal Medicine*. Philadelphia, PA: Elsevier (2008). p. 2–7. doi: 10.1016/j.siny.2007.09.005
19. Caughey AB, Cahill AG, Guise JM, Rouse DJ, of Obstetricians AC, Gynecologists, et al. Safe prevention of the primary cesarean delivery. *Am J Obstetr Gynecol*. (2014) 210:179–93. doi: 10.1016/j.ajog.2014.01.026
20. Bernardes J, Costa-Pereira A, Ayres-de Campos D, Van Geijn H, Pereira-Leite L. Evaluation of interobserver agreement of cardiotocograms. *Int J Gynecol Obstetr*. (1997) 57:33–7. doi: 10.1016/S0020-7292(97)02846-4
21. Donker DK, van Geijn HP, Hasman A. Interobserver variation in the assessment of fetal heart rate recordings. *Eur J Obstetr Gynecol Reproduct Biol*. (1993) 52:21–8. doi: 10.1016/0028-2243(93)90220-7
22. Huch A, Huch R, Rooth G. Guidelines for blood sampling and measurement of pH and blood gas values in obstetrics: based upon a workshop held in Zurich, Switzerland, March 19, 1993 by an Ad Hoc Committee. *Eur J Obstetr Gynecol Reproduct Biol*. (1994) 54:165–75. doi: 10.1016/0028-2243(94)90277-1
23. Ayres-de Campos D, Spong CY, Chandraharan E. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynecol Obstetr*. (2015) 131:13–24. doi: 10.1016/j.ijgo.2015.06.020
24. Schiermeier S, Pildner von Steinburg S, Thieme A, Reinhard J, Daumer M, Scholz M, et al. Sensitivity and specificity of intrapartum computerised FIGO criteria for cardiotocography and fetal scalp pH during labour: multicentre, observational study. *BJOG Int J Obstetr Gynaecol*. (2008) 115:1557–63. doi: 10.1111/j.1471-0528.2008.01857.x
25. Plsek PE, Greenhalgh T. The challenge of complexity in health care. *BMJ*. (2001) 323:625–8. doi: 10.1136/bmj.323.7313.625
26. Rotmensch S, Liberati M, Vishne T, Celentano C, Ben-Rafael Z, Bellati U. The effect of betamethasone and dexamethasone on fetal heart rate patterns and biophysical activities, a prospective randomized trial. *Acta Obstetr Gynecol Scand*. (1999) 78:493–500. doi: 10.1034/j.1600-0412.1999.780604.x
27. Signorini MG, Fanelli A, Magenes G. Monitoring fetal heart rate during pregnancy: contributions from advanced signal processing and wearable technology. *Comput Math Methods Med*. (2014) 2014: 707581. doi: 10.1155/2014/707581
28. Nunes I, Ayres-de Campos D, Figueiredo C, Bernardes J. An overview of central fetal monitoring systems in labour. *J Perinatal Med*. (2013) 41:93–9. doi: 10.1515/jpm-2012-0067
29. Wilson T, Holt T, Greenhalgh T. Complexity and clinical care. *BMJ*. (2001) 323:685–8. doi: 10.1136/bmj.323.7314.685
30. Chudáček V, Spilka J, Huptych M, Georgoulas G, Janků P P, Koucký M, et al. Automatic classification of intrapartal fetal heart-rate recordings-can it compete with experts? In: *International Conference on Information Technology in Bio Medical Informatics*. Berlin: Springer (2010) p. 57–66. doi: 10.1007/978-3-642-15020-3_5
31. Schnettler WT, Goldberger AL, Ralston SJ, Costa M. Complexity analysis of fetal heart rate preceding intrauterine demise. *Eur J Obstetr Gynecol Reproduct Biol*. (2016) 203:286–90. doi: 10.1016/j.ejogrb.2016.06.025
32. Chaffin DG, Goldberg CC, Reed KL. The dimension of chaos in the fetal heart rate. *Am J Obstetr Gynecol*. (1991) 165:1425–9. doi: 10.1016/S0002-9378(12)90778-4
33. Pincus SM, Viscarello RR. Approximate entropy: a regularity measure for fetal heart rate analysis. *Obstet Gynecol*. (1992) 79:249–55.
34. Datian Y, Xuemei O. Application of wavelet analysis in detection of fetal ECG. In: *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Amsterdam: IEEE (1996). p. 1043–4.
35. Signorini MG, Ferrario M, Pedrinazzi L, Magenes G. Analysis of echographic and heart rate time and frequency domain parameters for the antepartum fetal surveillance. In: *Computers in Cardiology*. Lyon: IEEE (2005). p. 679–82. doi: 10.1109/CIC.2005.1588193
36. Ferrario M, Signorini MG, Magenes G. Comparison between fetal heart rate standard parameters and complexity indexes for the identification of severe intrauterine growth restriction. *Methods Inform Med*. (2007) 46:186–90. doi: 10.1055/s-0038-1625404
37. Santos CC, Bernardes J, Vitányi PM, Antunes L. Clustering fetal heart rate tracings by compression. In: *19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)*. Salt Lake City, UT: IEEE (2006). p. 685–90.
38. Barquero-Pérez Ó, Santiago-Mozos R, Lillo-Castellano JM, Garcia-Viruete B, Goya-Esteban R, Caamano AJ, et al. Fetal heart rate analysis for automatic detection of perinatal hypoxia using normalized compression distance and machine learning. *Front Physiol*. (2017) 8:113. doi: 10.3389/fphys.2017.00113
39. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. (2009) 62:e1–34. doi: 10.1016/j.jclinepi.2009.06.006

40. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. (1948) 27:379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
41. Hoyer D, Nowack S, Bauer S, Tetschke F, Ludwig S, Moraru L, et al. Fetal development assessed by heart rate patterns-time scales of complex autonomic control. *Comput Biol Med*. (2012) 42:335–41. doi: 10.1016/j.compbiomed.2011.05.003
42. Frank B, Pompe B, Schneider U, Hoyer D. Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses. *Med Biol Eng Comput*. (2006) 44:179. doi: 10.1007/s11517-005-0015-z
43. Fulcher B, Georgieva A, Redman C. Highly comparative fetal heart rate analysis. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology-Proceedings*. San Diego, CA (2012). 3135–8. doi: 10.1109/EMBC.2012.6346629
44. Barquero-Pérez Ó, Goya-Esteban R, Caamano A, Martin-Caballero C, Rojo-Álvarez JL. Fetal heart rate complexity measures to detect hypoxia. In: *2015 Computing in Cardiology Conference (CinC)*. IEEE (2015). p. 133–6. doi: 10.1109/CIC.2015.7408604
45. Schneider U, Frank B, Fiedler A, Kaehler C, Hoyer D, Liehr M, et al. Human fetal heart rate variability-characteristics of autonomic regulation in the third trimester of gestation. *J Perinatal Med*. (2008) 36:433–41. doi: 10.1515/JPM.2008.059
46. Khandoker AH, Marzbanrad F, Voss A, Schulz S, Kimura Y, Endo M, et al. Analysis of maternal-fetal heart rate coupling directions with partial directed coherence. *Biomed Signal Process Control*. (2016) 30:25–30. doi: 10.1016/j.bspc.2016.06.010
47. Manis G, Sassi R. Relation between fetal HRV and value of umbilical cord artery pH in labor, a study with entropy measures. In: *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE (2017). p. 272–7. doi: 10.1109/CBMS.2017.139
48. Granero-Belinchon C, Roux SG, Abry P, Doret M, Garnier NB. Information theory to probe intrapartum fetal heart rate dynamics. *Entropy*. (2017) 19:640. doi: 10.3390/e19120640
49. Tetschke F, Schneider U, Schleussner E, Witte OW, Hoyer D. Assessment of fetal maturation age by heart rate variability measures using random forest methodology. *Comput Biol Med*. (2016) 70:157–62. doi: 10.1016/j.compbiomed.2016.01.020
50. Schneider U, Fiedler A, Schröder B, Jaekel S, Stacke A, Hoyer D, et al. The effect of antenatal steroid treatment on fetal autonomic heart rate regulation revealed by fetal magnetocardiography (fMCG). *Early Hum Dev*. (2010) 86:319–25. doi: 10.1016/j.earlhumdev.2010.05.018
51. Fergus P, Hussain A, Al-Jumeily D, Huang DS, Bouguila N. Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms. *Biomed Eng Online*. (2017) 16:89. doi: 10.1186/s12938-017-0378-z
52. Gonçalves H, Fernandes D, Pinto P, Ayres-de Campos D, Bernardes J. Simultaneous monitoring of maternal and fetal heart rate variability during labor in relation with fetal gender. *Dev Psychobiol*. (2017) 59:832–39. doi: 10.1002/dev.21554
53. Khandoker A, Karmakar C, Kimura Y, Endo M, Oshio S, Palaniswami M. Tone entropy analysis of foetal heart rate variability. *Entropy*. (2015) 17:1042–53. doi: 10.3390/e17031042
54. Hopkins P, Outram N, Lofgren N, Ifeachor EC, Rosén KG. A comparative study of fetal heart rate variability analysis techniques. In: *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE (2006). p. 1784–7. doi: 10.1109/IEMBS.2006.260258
55. Pincus SM, Gladstone IM, Ehrenkranz RA. A regularity statistic for medical data analysis. *J Clin Monitor*. (1991) 7:335–45. doi: 10.1007/BF01619355
56. Buchman TG. The community of the self. *Nature*. (2002) 420:246–51. doi: 10.1038/nature01260
57. Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys RevA*. (1992) 45:3403. doi: 10.1103/PhysRevA.45.3403
58. Cao L. Practical method for determining the minimum embedding dimension of a scalar time series. *Phys D Nonlinear Phenomena*. (1997) 110:43–50. doi: 10.1016/S0167-2789(97)00118-8
59. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Phys Rev A*. (1986) 33:1134. doi: 10.1103/PhysRevA.33.1134
60. Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci USA*. (1991) 88:2297–301. doi: 10.1073/pnas.88.6.2297
61. Pincus SM, Goldberger AL. Physiological time-series analysis: what does regularity quantify? *Am J Physiol*. (1994) 266:H1643. doi: 10.1152/ajpheart.1994.266.4.H1643
62. Marques de Sá JP. Characterization of fetal heart rate using approximate entropy. In: *Computers in Cardiology*. Lyon: IEEE (2005). p. 671–3. doi: 10.1109/CIC.2005.1588190
63. Magalhaes F, Marques de Sá J, Bernardes J, Ayres-de Campos D. Characterization of fetal heart rate irregularity using approximate entropy and wavelet filtering. In: *2006 Computers in Cardiology*. Valencia: IEEE (2006). p. 933–6.
64. Lu S, Chen X, Kanters JK, Solomon IC, Chon KH. Automatic selection of the threshold value for approximate entropy. *IEEE Trans Biomed Eng*. (2008) 55:1966–72. doi: 10.1109/TBME.2008.919870
65. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. (2000) 278:H2039–49. doi: 10.1152/ajpheart.2000.278.6.H2039
66. Costa M, Goldberger A, Peng CK. Multiscale entropy to distinguish physiologic and synthetic RR time series. In: *Computers in Cardiology*. Memphis, TN: IEEE (2002). p. 137–40.
67. Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of biological signals. *Phys Rev E*. (2005) 71:021906. doi: 10.1103/PhysRevE.71.021906
68. Valencia JF, Porta A, Vallverdu M, Claria F, Baranowski R, Orlowska-Baranowska E, et al. Refined multiscale entropy: application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects. *IEEE Trans Biomed Eng*. (2009) 56:2202–13. doi: 10.1109/TBME.2009.2021986
69. Nikulin VV, Brismar T. Comment on "Multiscale entropy analysis of complex physiologic time series". *Phys Rev Lett*. (2004) 92:089803. doi: 10.1103/PhysRevLett.92.089803
70. Lake DE, Moorman JR. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am J Physiol Heart Circ Physiol*. (2011) 300:H319–25. doi: 10.1152/ajpheart.00561.2010
71. Gierałtowski J, Hoyer D, Tetschke F, Nowack S, Schneider U, Zebrowski J. Development of multiscale complexity and multifractality of fetal heart rate variability. *Auton Neurosci*. (2013) 178:29–36. doi: 10.1016/j.autneu.2013.01.009
72. Kolmogorov AN. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. In: *Doklady Akademii Nauk*. Moscow: Russian Academy of Sciences (1958). p. 861–4.
73. Solomonoff RJ. A formal theory of inductive inference. Part II. *Inform Control*. (1964) 7:224–54. doi: 10.1016/S0019-9958(64)90131-7
74. Kolmogorov AN. Three approaches to the definition of the concept "quantity of information". *Problemy Peredachi Informatsii*. (1965) 1:3–11.
75. Chaitin GJ. On the length of programs for computing finite binary sequences: statistical considerations. *J ACM*. (1969) 16:145–59. doi: 10.1145/321495.321506
76. Lempel A, Ziv J. On the complexity of finite sequences. *IEEE Trans Inform Theory*. (1976) 22:75–81. doi: 10.1109/TIT.1976.1055501
77. Seward J. *bzip2*. (1996). Available online at: http://wwwmuraroademoncouk/
78. Spilka J, Chudáček V, Koucký M, Lhotská L, Huptych M, Janků P, et al. Using nonlinear features for fetal heart rate classification. *Biomed Signal Process Control*. (2012) 7:350–7. doi: 10.1016/j.bspc.2011.06.008
79. Yang ACC, Hseu SS, Yien HW, Goldberger AL, Peng CK. Linguistic analysis of the human heartbeat using frequency and rank order statistics. *Phys Rev Lett*. (2003) 90:108103. doi: 10.1103/PhysRevLett.90.108103
80. Ferrario M, Signorini MG, Magenes G. Complexity analysis of the fetal heart rate variability: early identification of severe intrauterine growth-restricted fetuses. *Med Biol Eng Comput*. (2009) 47:911–9. doi: 10.1007/s11517-009-0502-8
81. Magenes G, Bellazzi R, Fanelli A, Signorini MG. Multivariate analysis based on linear and non-linear FHR parameters for the identification of IUGR fetuses. In: *2014 36th Annual International Conference of the IEEE*

Engineering in Medicine and Biology Society. Chicago, IL: IEEE (2014). p. 1868–71. doi: 10.1109/EMBC.2014.6943974
82. Cilibrasi R, Vitányi PM. Clustering by compression. *IEEE Trans Inform Theory*. (2005) 51:1523–45. doi: 10.1109/TIT.2005.844059
83. Cilibrasi R, Vitányi P, Wolf Rd. Algorithmic clustering of music based on string compression. *Comput Music J*. (2004). 28:49–67. doi: 10.1162/0148926042728449
84. Wehner S. Analyzing worms and network traffic using compression. *J Comput Security*. (2007) 15:303–20. doi: 10.3233/JCS-2007-15301
85. Ferrario M, Signorini MG, Cerutti S. Complexity analysis of 24 hours heart rate variability time series. In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. San Francisco, CA: IEEE (2004). p. 3956–9.
86. Radhakrishnan N, Gangadhar B. Estimating regularity in epileptic seizure time-series data. *IEEE Eng Med Biol Mag*. (1998) 17:89–94. doi: 10.1109/51.677174
87. Zhang XS, Zhu YS, Thakor NV, Wang ZZ. Detecting ventricular tachycardia and fibrillation by complexity measure. *IEEE Trans Biomed Eng*. (1999) 46:548–55. doi: 10.1109/10.759055
88. Zhang XS, Roy RJ, Jensen EW. EEG complexity as a measure of depth of anesthesia for patients. *IEEE Trans Biomed Eng*. (2001) 48:1424–33. doi: 10.1109/10.966601
89. Szczepański J, Amigó JM, Wajnryb E, Sanchez-Vives M. Application of Lempel-Ziv complexity to the analysis of neural discharges. *Network Comput Neural Syst*. (2003) 14:335–50. doi: 10.1088/0954-898X_14_2_309
90. Castro L, Teixeira A, Brás S, Santos M, Costa-Santos C. Towards FHR biometric identification: a comparison between compression and entropy based approaches. In: *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. Karlstad: IEEE (2018). p. 440–1. doi: 10.1109/CBMS.2018.00085
91. Orlov YL, Potapov VN. Complexity: an internet resource for analysis of DNA sequence complexity. *Nucleic Acids Res*. (2004) 32:W628–33. doi: 10.1093/nar/gkh466
92. Katz MJ. Fractals and the analysis of waveforms. *Comput Biol Med*. (1988) 18:145–56. doi: 10.1016/0010-4825(88)90041-8
93. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Phys D Nonlinear Phenomena*. (1988) 31:277–83. doi: 10.1016/0167-2789(88)90081-4
94. Higuchi T. Relationship between the fractal dimension and the power law index for a time series: a numerical investigation. *Phys D Nonlinear Phenomena*. (1990) 46:254–64. doi: 10.1016/0167-2789(90)90039-R
95. Barabási AL. *Fractal Concepts in Surface Growth*. Cambridge, UK: Cambridge University Press (1995).
96. Henriques T, Ribeiro M, Teixeira A, Castro L, Antunes L, Costa-Santos C. Nonlinear methods most applied to heart-rate time series: a review. *Entropy*. (2020) 22:309. doi: 10.3390/e22030309
97. Hurst HE. Long-term storage capacity of reservoirs. *Trans Am Soc Eng*. (1951) 116:770–808. doi: 10.1061/TACEAT.0006518
98. Mandelbrot BB, Wallis JR. Noah, Joseph, and operational hydrology. *Water Resour Res*. (1968) 4:909–18. doi: 10.1029/WR004i005p00909
99. Mandelbrot BB, Wallis JR. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resour Res*. (1969) 5:967–88. doi: 10.1029/WR005i005p00967
100. Moreira J, da Silva JKL, Kamphorst SO. On the fractal dimension of self-affine profiles. *J Phys A Math Gen*. (1994) 27:8079. doi: 10.1088/0305-4470/27/24/018
101. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of DNA nucleotides. *Phys Rev E*. (1994) 49:1685. doi: 10.1103/PhysRevE.49.1685
102. Geweke J, Porter-Hudak S. The estimation and application of long memory time series models. *J Time Ser Anal*. (1983) 4:221–38. doi: 10.1111/j.1467-9892.1983.tb00371.x
103. Beran J. *Statistics for Long-Memory Processes*. vol. 61. Boca Raton, FL: CRC Press (1994).
104. Robinson PM, et al. Gaussian semiparametric estimation of long range dependence. *Ann Stat*. (1995) 23:1630–61. doi: 10.1214/aos/1176324317
105. Hansen A, Engøy T, Måløy KJ. Measuring Hurst exponents with the first return method. *Fractals*. (1994) 2:527–33. doi: 10.1142/S0218348X94000740
106. Abry P, Gonçalvés P, Flandrin P. Wavelets, spectrum analysis and 1/f processes. In: Antoniadis A, Oppenheim G, editors, *Wavelets and Statistics Lecture Notes in Statistics*, vol 103. New York, NY: Springer (1995). p. 15–29. doi: 10.1007/978-1-4612-2544-7_2
107. Schepers HE, Van Beek JH, Bassingthwaighte JB. Four methods to estimate the fractal dimension from self-affine signals (medical application). In: *Antoniadis A, Oppenheim G, editors. IEEE Engineering in Medicine and Biology Magazine*. New York, NY: Springer (1992). 11:57–64. doi: 10.1109/51.139038
108. Peng CK, Havlin S, Hausdorff JM, Mietus JE, Stanley HE, Goldberger AL. Fractal mechanisms and heart rate dynamics: long-range correlations and their breakdown with disease. *J Electrocardiol*. (1995) 28:59–65. doi: 10.1016/S0022-0736(95)80017-4
109. Kantelhardt JW, Koscielny-Bunde E, Rego HHA, Havlin S, Bunde A. Detecting long-range correlations with detrended fluctuation analysis. *Phys A Stat Mech Appl*. (2001) 295:441–54. doi: 10.1016/S0378-4371(01)00144-3
110. Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. *Phys A Stat Mech Appl*. (2002) 316:87–114. doi: 10.1016/S0378-4371(02)01383-3
111. Ossadnik S, Buldyrev SV, Goldberger AL, Havlin S, Mantegna RN, Peng CK, et al. Correlation approach to identify coding regions in DNA sequences. *Biophys J*. (1994) 67:64–70. doi: 10.1016/S0006-3495(94)80455-2
112. Bunde A, Havlin S, Kantelhardt JW, Penzel T, Peter JH, Voigt K. Correlated and uncorrelated regions in heart-rate fluctuations during sleep. *Phys Rev Lett*. (2000) 85:3736. doi: 10.1103/PhysRevLett.85.3736
113. Gierałtowski J, Zebrowski J, Baranowski R. Multiscale multifractal analysis of heart rate variability recordings with a large number of occurrences of arrhythmia. *Phys Rev E*. (2012) 85:021915. doi: 10.1103/PhysRevE.85.021915
114. Haar A. *Zur Theorie der Orthogonalen Funktionensysteme*. Georg-August-Universitat, Gottingen (1909). doi: 10.1007/BF01456326
115. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell*. (1989) 11:674–93. doi: 10.1109/34.192463
116. Daubechies I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans Inform Theory*. (1990) 36:961–1005. doi: 10.1109/18.57199
117. Daubechies I, Bates BJ. *Ten Lectures on Wavelets*. Philadelphia, PA: Acoustical Society of America (1993). doi: 10.1137/1.9781611970104
118. Mallat S. *A Wavelet Tour of Signal Processing*. Los Angeles, CA: Elsevier (1999). doi: 10.1016/B978-012466606-1/50008-8
119. Desai K, Jadhav SD, Sankhe MS. A comparison and quantification of fetal heart rate variability using Doppler ultrasound and direct electrocardiography acquisition techniques. In: *2013 International Conference on Advances in Technology and Engineering (ICATE)*. Mumbai: IEEE (2013). p. 1–8. doi: 10.1109/ICAdTE.2013.6524738
120. Gilles J. Empirical wavelet transform. *IEEE Trans Signal Process*. (2013) 61:3999–4010. doi: 10.1109/TSP.2013.2265222
121. Zhao Z, Deng Y, Zhang Y, Zhang Y, Zhang X, Shao L. DeepFHR: intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network. *BMC Med Inform Decis Mak*. (2019) 19:286. doi: 10.1186/s12911-019-1007-5
122. Papadimitriou S, Papadopoulos V, Gatzounas D, Tzigounis V, Bezerianos A. The performance and reliability of wavelet denoising for Doppler ultrasound fetal heart rate signal preprocessing. *Stud Health Technol Inform*. (1997) 43:561–5.
123. Barnova K, Martinek R, Jaros R, Kahankova R. Hybrid methods based on empirical mode decomposition for non-invasive fetal heart rate monitoring. *IEEE Access*. (2020) 8:51200–18. doi: 10.1109/ACCESS.2020.2980254
124. Meddour C, Kedir-Talha M. New method exploiting a hybrid techniques for fetal cardiac signal extraction. *Biomed Eng Appl Basis Commun*. (2019) 31:1950027. doi: 10.4015/S1016237219500273
125. Jaros R, Martinek R, Kahankova R, Koziorek J. Novel hybrid extraction systems for fetal heart rate variability monitoring based on non-invasive fetal electrocardiogram. *IEEE Access*. (2019) 7:131758–131784. doi: 10.1109/ACCESS.2019.2933717

126. Georgoulas G, Stylios C, Groumpos P. Feature extraction and classification of fetal heart rate using wavelet analysis and support vector machines. *Int J Artif Intell Tools*. (2006) 15:411–32. doi: 10.1142/S0218213006002746
127. 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 Obstetr Gynecol*. (2010) 202:258. doi: 10.1016/j.ajog.2009.06.026
128. Cömert Z, Şengür A, Budak Ü, Kocamaz AF. Prediction of intrapartum fetal hypoxia considering feature selection algorithms and machine learning models. *Health Inform Sci Syst*. (2019) 7:1–9. doi: 10.1007/s13755-019-0079-z
129. Cömert Z, Kocamaz AF. Open-access software for analysis of fetal heart rate signals. *Biomed Signal Process Control*. (2018) 45:98–108. doi: 10.1016/j.bspc.2018.05.016
130. Monteiro-Santos J, Gonçalves H, Bernardes J, Antunes L, Nozari M, Costa-Santos C. Entropy and compression capture different complexity features: the case of fetal heart rate. *Entropy*. (2017) 19:688. doi: 10.3390/e19120688
131. Gonçalves H, Ayres-de Campos D, Bernardes J. The effect of gender, gestational age and behavioral states on fetal heart rate variability. In: *2014 8th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*. IEEE (2014). p. 19–20. doi: 10.1109/ESGCO.2014.6847498
132. Gonçalves H, Costa A, Ayres-de Campos D, Costa-Santos C, Rocha AP, Bernardes J. Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability. *Med Biol Eng Comput*. (2013) 51:665–76. doi: 10.1007/s11517-013-1036-7
133. Bernardes J, Gonçalves H, Ayres-de Campos D, Rocha AP. Sex differences in linear and complex fetal heart rate dynamics of normal and acidemic fetuses in the minutes preceding delivery. *J Perinatal Med*. (2009) 37:168–76. doi: 10.1515/JPM.2009.024
134. Salamalekis E, Thomopoulos P, Giannaris D, Salloum I, Vasios G, Prentza A, et al. Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilising wavelet analysis and neural networks. *BJOG Int J Obstetr Gynaecol*. (2002) 109:1137–42. doi: 10.1111/j.1471-0528.2002.01388.x
135. Vasios G, Prentza A, Blana D, Salamalekis E, Thomopoulos P, Giannaris D, et al. Classification of fetal heart rate tracings based on wavelet-transformand self-organizing-map neural networks. In: *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Istanbul: IEEE (2001). p. 1633–6.
136. Kimura Y, Okamura K, Watanabe T, Yaegashi N, Uehara S, Yajima A. Time-frequency analysis of fetal heartbeat fluctuation using wavelet transform. *Am J Physiol Heart Circ Physiol*. (1998) 275:H1993–9. doi: 10.1152/ajpheart.1998.275.6.H1993
137. Zhao Z, Zhang Y, Comert Z, Deng Y. Computer-aided diagnosis system of fetal hypoxia incorporating recurrence plot with convolutional neural network. *Front Physiol*. (2019) 10:255. doi: 10.3389/fphys.2019.00255
138. Zhao Z, Zhang Y, Deng Y. A comprehensive feature analysis of the fetal heart rate signal for the intelligent assessment of fetal state. *J Clin Med*. (2018) 7:223. doi: 10.3390/jcm7080223
139. Cömert Z, Kocamaz AF, Subha V. Prognostic model based on image-based time-frequency features and genetic algorithm for fetal hypoxia assessment. *Comput Biol Med*. (2018) 99:85–97. doi: 10.1016/j.compbiomed.2018.06.003
140. Arican M, Cömer Z, Kocamaz AF, Polat K. Analysis of fetal heart rate signal based on neighborhood-based variance compression method. In: *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*. Malatya: IEEE (2018). p. 1–6. doi: 10.1109/IDAP.2018.8620898
141. Gonçalves H, Pinto P, Silva M, Ayres-de Campos D, Bernardes J. Toward the improvement in fetal monitoring during labor with the inclusion of maternal heart rate analysis. *Med Biol Eng Comput*. (2016) 54:691–9. doi: 10.1007/s11517-015-1359-7
142. Pasarica A, Rotariu C, Bozomitu RG, Eva OD. Dynamic of couplings between fetal heart rate and uterine contractions. In: *2015 International Symposium on Signals, Circuits and Systems (ISSCS)*. Iasi: IEEE (2015). p. 1–4. doi: 10.1109/ISSCS.2015.7203968
143. Chudáček V, Spilka J, Jankŭ P, Koucký M, Lhotská L, Huptych M. Automatic evaluation of intrapartum fetal heart rate recordings: a comprehensive analysis of useful features. *Physiol Measure*. (2011) 32:1347. doi: 10.1088/0967-3334/32/8/022
144. Gonçalves H, Rocha AP, Ayres-de Campos D, Bernardes J. Frequency domain and entropy analysis of fetal heart rate: appealing tools for fetal surveillance and pharmacodynamic assessment of drugs. *Cardiovasc Haematol Disord Drug Targets*. (2008) 8:91–8. doi: 10.2174/187152908784533720
145. Henriques T, Gonçalves H, Antunes L, Matias M, Bernardes J, Costa-Santos C. Entropy and compression: two measures of complexity. *J Eval Clin Pract*. (2013) 19:1101–6. doi: 10.1111/jep.12068
146. Gavrilis D, Tsoulos I. Classification of fetal heart rate using grammatical evolution. In: *IEEE Workshop on Signal Processing Systems Design and Implementation*. (2005). IEEE (2005). p. 425–9.
147. Fuentealba P, Illanes A, Ortmeier F. Cardiotocographic signal feature extraction through ceemdan and time-varying autoregressive spectral-based analysis for fetal welfare assessment. *IEEE Access*. (2019) 7:159754–72. doi: 10.1109/ACCESS.2019.2950798
148. Granero-Belinchon C, Roux SG, Garnier NB, Abry P, Doret M. Mutual information for intrapartum fetal heart rate analysis. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Jeju: IEEE (2017). p. 2014–7. doi: 10.1109/EMBC.2017.8037247
149. Georgoulas G, Karvelis P, Spilka J, Chudáček V, Stylios CD, Lhotska L. Investigating pH based evaluation of fetal heart rate (FHR) recordings. *Health Technol*. (2017) 7:241–54. doi: 10.1007/s12553-017-0201-7
150. Stylios CD, Georgoulas G, Karvelis P, Spilka J, Chudáček V, Lhotska L. Least squares support vector machines for FHR classification and assessing the pH based categorization. In: *XIV Mediterranean Conference on Medical and Biological Engineering and Computing*. Paphos: Springer (2016). p. 1211–5. doi: 10.1007/978-3-319-32703-7_234
151. Warmerdam G, Vullings R, Van Laar JO, Bergmans J, Schmitt L, Oei S, et al. Using uterine activity to improve fetal heart rate variability analysis for detection of asphyxia during labor. *Physiol Measure*. (2016) 37:387. doi: 10.1088/0967-3334/37/3/387
152. Spilka J, Frecon J, Leonarduzzi R, Pustelnik N, Abry P, Doret M. Intrapartum fetal heart rate classification from trajectory in sparse SVM feature space. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan: IEEE (2015). p. 2335–8. doi: 10.1109/EMBC.2015.7318861
153. Spilka J, Abry P, Goncalves P, Doret M. Impacts of first and second labour stages on Hurst parameter based intrapartum fetal heart rate analysis. In: *Computing in Cardiology*. Cambridge, MA: IEEE (2014). p. 777–780.
154. 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:e0136661. doi: 10.1371/journal.pone.0136661
155. Leonarduzzi R, Spilka J, Wendt H, Jaffard S, Torres ME, Abry P, et al. p-leader based classification of first stage intrapartum fetal HRV. In: *VI Latin American Congress on Biomedical Engineering CLAIB*. Paraná: Springer (2015). p. 504–7. doi: 10.1007/978-3-319-13117-7_129
156. 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*. (2013) 61:1100–8. doi: 10.1109/TBME.2013.2294324
157. Abry P, Roux SG, Chudáček V, Borgnat P, Gonçalves P, Doret M. Hurst exponent and intrapartum fetal heart rate: impact of decelerations. In: *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems. Porto: IEEE (2013)* p. 131–6. doi: 10.1109/CBMS.2013.6627777
158. Spilka J, Roux SG, Garnier NB, Abry P, Gonçalves P, Doret M. Nearestneighbor based wavelet entropy rate measures for intrapartum fetal heart rate variability. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Chicago, IL: IEEE (2014). p. 2813–6. doi: 10.1109/EMBC.2014.6944208
159. Costa MD, Schnettler WT, Amorim-Costa C, Bernardes J, Costa A, Goldberger AL, et al. Complexity-loss in fetal heart rate dynamics during labor as a potential biomarker of acidemia. *Early Hum Dev*. (2014) 90:67–71. doi: 10.1016/j.earlhumdev.2013.10.002
160. Santiago-Mozos R, Garcia-Vizuete B, Lillo-Castellano JM, Rojo-Alvarez JL, Martin-Caballero C. On the early detection of perinatal hypoxia with

information-theory based methods. In: *Computing in Cardiology*. Zaragoza: IEEE (2013). p. 425–8.
161. Doret M, Helgason H, Abry P, Goncalves P, Gharib C, Gaucherand P. Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor. *Am J Perinatol*. (2011) 28:259–66. doi: 10.1055/s-0030-1268713
162. Abry P, Helgason H, Gonçalves P, Pereira E, Gaucherand P, Doret M. Multifractal analysis of ECG for intrapartum diagnosis of fetal asphyxia. In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. Dallas, TX: IEEE (2010). p. 566–9. doi: 10.1109/ICASSP.2010.5495583
163. Spilka J, Chudáček V, Koucký M, Lhotská L. Assessment of non-linear features for intrapartal fetal heart rate classification. In: *2009 9th International Conference on Information Technology and Applications in Biomedicine*. Larnaka: IEEE (2009). p. 1–4. doi: 10.1109/ITAB.2009.5394442
164. Gonçalves H, Rocha AP, Ayres-de Campos D, Bernardes J. Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery. *Med Biol Eng Comput*. (2006) 44:847. doi: 10.1007/s11517-006-0105-6
165. Salamalekis E, Siristatidis C, Vasios G, Saloum J, Giannaris D, Chrelias C, et al. Fetal pulse oximetry and wavelet analysis of the fetal heart rate in the evaluation of abnormal cardiotocography tracings. *J Obstetr Gynaecol Res*. (2006) 32:135–139. doi: 10.1111/j.1447-0756.2006.00377.x
166. Cömert Z, Kocamaz AF. A novel software for comprehensive analysis of cardiotocography signals "CTG-OAS". In: *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*. Malatya: IEEE (2017). p. 1–6. doi: 10.1109/IDAP.2017.8090210
167. Chudáček V, Spilka J, Lhotská L, Janků P, Koucký M, Huptych M, et al. Assessment of features for automatic CTG analysis based on expert annotation. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Boston, MA: IEEE (2011). p. 6051–4.
168. d'Aloja E, Müller M, Paribello F, Demontis R, Faa A. Neonatal asphyxia and forensic medicine. *J Maternal Fetal Neonatal Med*. (2009) 22:54–6. doi: 10.1080/14767050903198397
169. Heintz E, Brodtkorb TH, Nelson N, Levin LÅ. The long-term cost-effectiveness of fetal monitoring during labour: a comparison of cardiotocography complemented with ST analysis versus cardiotocography alone. *BJOG Int J Obstetr Gynaecol*. (2008) 115:1676–87. doi: 10.1111/j.1471-0528.2008.01935.x
170. Malcus P. Antenatal fetal surveillance. *Curr Opin Obstetr Gynecol*. (2004) 16:123–8. doi: 10.1097/00001703-200404000-00005
171. Florio P, Marinoni E, Di Iorio R, Bashir M, Ciotti S, Sacchi R, et al. Urinary S100B protein concentrations are increased in intrauterine growth-retarded newborns. *Pediatrics*. (2006) 118:e747–54. doi: 10.1542/peds.2005-2875
172. Zeitlin J, El Ayoubi M, Jarreau PH, Draper ES, Blondel B, Künzel W, et al. Impact of fetal growth restriction on mortality and morbidity in a very preterm birth cohort. *J Pediatr*. (2010) 157:733–9. doi: 10.1016/j.jpeds.2010.05.002
173. Figueras F, Gardosi J. Intrauterine growth restriction: new concepts in antenatal surveillance, diagnosis, and management. *Am J Obstetr Gynecol*. (2011) 204:288–300. doi: 10.1016/j.ajog.2010.08.055
174. Fukushima T, Flores CA, Hon EH, Davidson EC Jr. Limitations of autocorrelation in fetal heart rate monitoring. *Am J Obstetr Gynecol*. (1985) 153:685–92. doi: 10.1016/S0002-9378(85)80261-1
175. Gonçalves H, Rocha AP, Ayres-de Campos D, Bernardes J. Internal versus external intrapartum foetal heart rate monitoring: the effect on linear and nonlinear parameters. *Physiol Measure*. (2006) 27:307. doi: 10.1088/0967-3334/27/3/008
176. Van Leeuwen P, Geue D, Thiel M, Cysarz D, Lange S, Romano MC, et al. Influence of paced maternal breathing on fetal-maternal heart rate coordination. *Proc Natl Acad Sci USA*. (2009) 106:13661–6. doi: 10.1073/pnas.0901049106
177. Spyridou K, Chouvarda I, Hadjileontiadis L, Maglaveras N. The effect of cigarette smoking on fetal heart rate tracing during pregnancy. *J Perinatal Med*. (2017) 45:403–11. doi: 10.1515/jpm-2015-0275
178. Van Leeuwen P, Cysarz D, Edelhäuser F, Grönemeyer D. Heart rate variability in the individual fetus. *Auton Neurosci*. (2013) 178:24–8. doi: 10.1016/j.autneu.2013.01.005
179. Montenegro N, Ramos C, Matias A, Barros H. Variation of embryonic/fetal heart rate at 6-13 weeks' gestation. *Ultrasound Obstetr Gynecol*. (1998) 11:274–6. doi: 10.1046/j.1469-0705.1998.11040274.x
180. Arduini D, Rizzo G, Caforio L, Boccolini MR, Romanini C, Mancuso S. Behavioural state transitions in healthy and growth retarded fetuses. *Early Hum Dev*. (1989) 19:155–65. doi: 10.1016/0378-3782(89)90076-5
181. DiPietro JA, Bornstein MH, Hahn CS, Costigan K, Achy-Brou A. Fetal heart rate and variability: stability and prediction to developmental outcomes in early childhood. *Child Dev*. (2007) 78:1788–98. doi: 10.1111/j.1467-8624.2007.01099.x
182. Spyridou K, Chouvarda I, Hadjileontiadis L, Maglaveras N. Linear and nonlinear features of fetal heart rate on the assessment of fetal development in the course of pregnancy and the impact of fetal gender. *Physiol Measure*. (2018) 39:015007. doi: 10.1088/1361-6579/aa9e3c
183. Tagliaferri S, Esposito FG, Fagioli R, Di Cresce M, Sacchi L, Signorini MG, et al. Ethnic analogies and differences in fetal heart rate variability signal: a retrospective study. *J Obstetr Gynaecol Res*. (2017) 43:281–290. doi: 10.1111/jog.13213
184. Tendais I, Figueiredo B, Gonçalves H, Bernardes J, Ayres-de Campos D, Montenegro N. Sex differences in the fetal heart rate variability indices of twins. *J Perinatal Med*. (2015) 43:221–5. doi: 10.1515/jpm-2014-0031
185. Park YS, Ryu KY, Shim SS, Hoh JK, Park MI. Comparison of fetal heart rate patterns using nonlinear dynamics in breech versus cephalic presentation at term. *Early Hum Dev*. (2013) 89:101–6. doi: 10.1016/j.earlhumdev.2012.08.006
186. Gonçalves H, Ayres-de Campos D, Bernardes J. Fetal behavioral dynamics in cephalic versus breech presentations. *Dev Psychobiol*. (2014) 56:1595–600. doi: 10.1002/dev.21242
187. Takashima T, Koyanagi T, Horimoto N, Satoh S, Nakano H. Breech presentation: is there a difference in eye movement patterns compared with cephalic presentation in the human fetus at term? *Am J Obstetr Gynecol*. (1995) 172:851–5. doi: 10.1016/0002-9378(95)90010-1
188. Kean L, Suwanrath C, Gargari S, Sahota D, James D. A comparison of fetal behaviour in breech and cephalic presentations at term. *BJOG Int J Obstetr Gynaecol*. (1999) 106:1209–13. doi: 10.1111/j.1471-0528.1999.tb08150.x
189. Choi WY, Hoh JK. Nonlinear analysis of fetal heart rate dynamics in fetuses compromised by asymptomatic partial placental abruption. *Placenta*. (2015) 36:1474–9. doi: 10.1016/j.placenta.2015.10.002
190. Moraes ER, Murta Jr LO, Baffa O, Wakai RT, Comani S. Linear and nonlinear measures of fetal heart rate patterns evaluated on very short fetal magnetocardiograms. *Physiol Measure*. (2012) 33:1563. doi: 10.1088/0967-3334/33/10/1563
191. Nijhuis J, Prechtl H, Martin CB Jr, Bots R. Are there behavioural states in the human fetus? *Early Hum Dev*. (1982) 6:177–95. doi: 10.1016/0378-3782(82)90106-2
192. Nijhuis IJ, ten Hof J. Development of fetal heart rate and behavior: indirect measures to assess the fetal nervous system. *Eur J Obstetr Gynecol*. (1999) 1:1–2. doi: 10.1016/S0301-2115(99)00143-8
193. Swartjes J, Van Geijn H, Mantel R, Van Woerden E, Schoemaker H. Coincidence of behavioural state parameters in the human fetus at three gestational ages. *Early Hum Dev*. (1990) 23:75–83. doi: 10.1016/0378-3782(90)90130-B
194. Jiménez-González A, James C. On the interpretation of the independent components underlying the abdominal phonogram: a study of their physiological relevance. *Physiol Measure*. (2012) 33:297. doi: 10.1088/0967-3334/33/2/297
195. Silva I, Behar J, Sameni R, Zhu T, Oster J, Clifford GD, et al. Noninvasive fetal ECG: the PhysioNet/computing in cardiology challenge. In: *Computing in Cardiology*. Zaragoza: IEEE (2013). p. 149–52.
196. Di Maria C, Liu C, Zheng D, Murray A, Langley P. Extracting fetal heart beats from maternal abdominal recordings: selection of the optimal principal components. *Physiol Measure*. (2014) 35:1649. doi: 10.1088/0967-3334/35/8/1649
197. Alnuaimi S, Jimaa S, Kimura Y, Apostolidis GK, Hadjileontiadis LJ, Khandoker AH. Fetal cardiac timing events estimation from Doppler

ultrasound signals using swarm decomposition. *Front Physiol*. (2019) 10:789. doi: 10.3389/fphys.2019.00789
198. Lin C, Yeh CH, Wang CY, Shi W, Serafico BMF, Wang CH, et al. Robust fetal heart beat detection via R-peak intervals distribution. *IEEE Trans Biomed Eng*. (2019) 66:3310–9. doi: 10.1109/TBME.2019.2904014
199. Monteiro-Santos J, Henriques T, Nunes I, Amorim-Costa C, Bernardes J, Costa-Santos C. Complexity of cardiotocographic signals as a predictor of labor. *Entropy*. (2020) 22:104. doi: 10.3390/e22010104
200. Kim KN, Park YS, Hoh JK. Sex-related differences in the development of fetal heart rate dynamics. *Early Hum Dev*. (2016) 93:47–55. doi: 10.1016/j.earlhumdev.2015.12.005
201. Montalvo-Jaramillo CI, Pliego-Carrillo AC, Pena-Castillo MÁ, Echeverria JC, Becerril-Villanueva E, Pavón L, et al. Comparison of fetal heart rate variability by symbolic dynamics at the third trimester of pregnancy and low-risk parturition. *Heliyon*. (2020) 6:e03485. doi: 10.1016/j.heliyon.2020.e03485
202. Lim J, Kwon JY, Song J, Choi H, Shin JC, Park IY. Quantitative comparison of entropy analysis of fetal heart rate variability related to the different stages of labor. *Early Hum Dev*. (2014) 90:81–85. doi: 10.1016/j.earlhumdev.2013.12.007
203. Saleem S, Naqvi SS, Manzoor T, Saeed A, Mirza J, et al. A strategy for classification of "vaginal vs. cesarean section" delivery: bivariate empirical mode decomposition of cardiotocographic recordings. *Front Physiol*. (2019) 10:246. doi: 10.3389/fphys.2019.00246
204. Gonçalves H, Amorim-Costa C, Ayres-de Campos D, Bernardes J. Evolution of linear and nonlinear fetal heart rate indices throughout pregnancy in appropriate, small for gestational age and preterm fetuses: a cohort study. *Comput Methods Prog Biomed*. (2018) 153:191–9. doi: 10.1016/j.cmpb.2017.10.015
205. Tagliaferri S, Esposito FG, Esposito G, Saccone G, Signorini MG, Magenes G, et al. Impact of nuchal cord on antenatal and intrapartum foetal heart rate surveillance and perinatal outcome. *J Obstetr Gynaecol*. (2020) 40:316–23. doi: 10.1080/01443615.2019.1621816
206. Boudet S, de l'Aulnoit AH, Demailly R, Peyrodie L, Beuscart R, de l'Aulnoit DH. Fetal heart rate baseline computation with a weighted median filter. *Comput Biol Med*. (2019) 114:103468. doi: 10.1016/j.compbiomed.2019.103468
207. Magenes G, Signorini MG, Arduini D, Cerutti S. Fetal heart rate variability due to vibroacoustic stimulation: linear and nonlinear contribution. *Methods Inform Med*. (2004) 43:47–51. doi: 10.1055/s-0038-1633833
208. Annunziata ML, Scala M, Giuliano N, Tagliaferri S, Imperato OCM, Esposito FG, et al. Fetal vibroacoustic stimulation in computerized cardiotocographic analysis: the role of short-term variability and approximate entropy. *J Pregnancy*. (2012) 2012:814987. doi: 10.1155/2012/814987
209. Signorini MG, Magenes G, Cerutti S, Arduini D. Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings. *IEEE Trans Biomed Eng*. (2003) 50:365–74. doi: 10.1109/TBME.2003.808824
210. Signorini MG, Pini N, Malovini A, Bellazzi R, Magenes G. Dataset on linear and non-linear indices for discriminating healthy and IUGR fetuses. *Data Brief*. (2020) 29:105164. doi: 10.1016/j.dib.2020.105164
211. Chen K, Zhao Y, Li S, Chen L, Wang N, Zhang K, et al. Multiscale coupling of uterine electromyography and fetal heart rate as a novel indicator of fetal neural development. *Front Neurol*. (2019) 10:760. doi: 10.3389/fneur.2019.00760
212. Esposito FG, Tagliaferri S, Giudicepietro A, Giuliano N, Maruotti GM, Saccone G, et al. Fetal heart rate monitoring and neonatal outcome in a population of early-and late-onset intrauterine growth restriction. *J Obstetr Gynaecol Res*. (2019) 45:1343–351. doi: 10.1111/jog.13981
213. Fuentealba P, Illanes A, Ortmeier F. Independent analysis of decelerations and resting periods through CEEMDAN and spectral-based feature extraction improves cardiotocographic assessment. *Appl Sci*. (2019) 9:5421. doi: 10.3390/app9245421
214. Fuentealba P, Illanes A, Ortmeier F. Cardiotocograph data classification improvement by using empirical mode decomposition. In: *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Berlin: IEEE (2019). p. 5646–9. doi: 10.1109/EMBC.2019.8856673
215. Frasch MG, Herry CL, Niu Y, Giussani DA. First evidence that intrinsic fetal heart rate variability exists and is affected by hypoxic pregnancy. *J Physiol*. (2020) 598:249–63. doi: 10.1113/JP278773
216. Khandoker AH, Schulz S, Al-Angari HM, Voss A, Kimura Y. Alterations in maternal-fetal heart rate coupling strength and directions in abnormal fetuses. *Front Physiol*. (2019) 10:482. doi: 10.3389/fphys.2019.00482
217. Marques JAL, Cortez PC, Madeiro JP, de Albuquerque VHC, Fong SJ, Schlindwein FS. Nonlinear characterization and complexity analysis of cardiotocographic examinations using entropy measures. *J Supercomput*. (2020) 76:1305–20. doi: 10.1007/s11227-018-2570-8
218. Keenan E, Udhayakumar RK, Karmakar CK, Brownfoot FC, Palaniswami M. Entropy profiling for detection of fetal arrhythmias in short length fetal heart rate recordings. In: *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. Montreal, QC: IEEE (2020). p. 621–4. doi: 10.1109/EMBC44109.2020.9175892
219. Hoyer D, Schmidt A, Gustafson KM, Lobmaier SM, Lakhno I, van Leeuwen P, et al. Heart rate variability categories of fluctuation amplitude and complexity: diagnostic markers of fetal development and its disturbances. *Physiol Measure*. (2019) 40:064002. doi: 10.1088/1361-6579/ab205f
220. Park YS, Hoh JK. Complex and irregular heart rate dynamics in fetuses compromised by maternal anemia as a high-risk pregnancy. *J Perinatal Med*. (2015) 43:741–748. doi: 10.1515/jpm-2014-0104
221. Alshebly Y, Nafea M. Isolation of fetal ECG signals from abdominal ECG using wavelet analysis. *IRBM*. (2020) 41:252–60. doi: 10.1016/j.irbm.2019.12.002
222. John RG, Ramachandran K. Extraction of foetal ECG from abdominal ECG by nonlinear transformation and estimations. *Comput Methods Prog Biomed*. (2019) 175:193–204. doi: 10.1016/j.cmpb.2019.04.022
223. Al-Sheikh B, Salman MS, Eleyan A, Alboon S. Non-invasive fetal ECG extraction using discrete wavelet transform recursive inverse adaptive algorithm. *Technol Health Care*. (2019) 28:507–20. doi: 10.3233/THC-191948
224. Yuan L, Yuan Y, Zhou Z, Bai Y, Wu S. A fetal ECG monitoring system based on the android smartphone. *Sensors*. (2019) 19:446. doi: 10.3390/s19030446
225. Ayres-de Campos D, Bernardes J, Garrido A, Marques de Sa J, Pereira-Leite L. SisPorto 2.0: a program for automated analysis of cardiotocograms. *J Maternal Fetal Med*. (2000) 9:311–8. doi: 10.1002/1520-6661(200009/10)9:5<311::AID-MFM12>3.0.CO;2-9
226. Georgieva A, Payne SJ, Moulden M, Redman CW. Artificial neural networks applied to fetal monitoring in labour. *Neural Comput Appl*. (2013) 22:85–93. doi: 10.1007/s00521-011-0743-y
227. Guijarro-Berdinas B, Alonso-Betanzos A, Fontenla-Romero O. Intelligent analysis and pattern recognition in cardiotocographic signals using a tightly coupled hybrid system. *Artif Intell*. (2002) 136:1–27. doi: 10.1016/S0004-3702(01)00163-1
228. Amer-Wåhlin I, Maršál K. ST analysis of fetal electrocardiography in labor. In: *Seminars in Fetal and Neonatal Medicine*. Philadelphia, PA: WB Saunders (2011). p. 29–35. doi: 10.1016/j.siny.2010.09.004
229. Ribeiro M, Henriques T, Castro L, Souto A, Antunes L, Costa-Santos C, et al. The entropy universe. *Entropy*. (2021) 23:222. doi: 10.3390/e23020222
230. Tomassini S, Strazza A, Sbrollini A, Marcantoni I, Morettini M, Fioretti S, et al. Wavelet filtering of fetal phonocardiography: a comparative analysis. *Math Biosci Eng*. (2019) 16:6034–46. doi: 10.3934/mbe.2019302
231. Dawes G, Moulden M, Redman C. Criteria for the design of fetal heart rate analysis systems. *Int J biomed Comput*. (1990) 25:287–94. doi: 10.1016/0020-7101(90)90032-P
232. Maulik D, Saini V, Zigrossi S. Clinical significance of short-term variability computed from heart-rate waveforms. *J Perinatal Med*. (1983) 11:243–8. doi: 10.1515/jpme.1983.11.5.243
233. Modanlou HD, Freeman RK, Braly P. A simple method of fetal and neonatal heart rate beat-to-beat variability quantitation: Preliminary report. *Am J Obstetr Gynecol*. (1977) 127:861–8. doi: 10.1016/0002-9378(77)90119-3
234. Druzen M, Ikenoye T, Murata Y. A possible mechanism for the increase in FHR variability following hypoxemia. In: *Presented before the 26th Annual Meeting of the Society for Gynecologic Investigation*. San Diego, CA (1979).

235. Ferrario M, Signorini MG, Magenes G, Cerutti S. Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress. *IEEE Trans Biomed Eng*. (2006) 53:119–25. doi: 10.1109/TBME.2005.859809
236. Baschat A. Neurodevelopment following fetal growth restriction and its relationship with antepartum parameters of placental dysfunction. *Ultrasound Obstetr Gynecol*. (2011) 37:501–14. doi: 10.1002/uog.9008
237. Tagliaferri S, Fanelli A, Esposito G, Esposito FG, Magenes G, Signorini MG, et al. Evaluation of the acceleration and deceleration phase-rectified slope to detect and improve IUGR clinical management. *Comput Math Methods Med*. (2015) 2015:236896. doi: 10.1155/2015/236896
238. Bernardes J, Moura C, Marques de Sa JP, Leite LP. The Porto system for automated cardiotocographic signal analysis. *J Perinatal Med*. (1991) 19:61–5. doi: 10.1515/jpme.1991.19.1-2.61
239. Ivanov PC, Ma QD, Bartsch RP. Maternal-fetal heartbeat phase synchronization. *Proc Natl Acad Sci USA*. (2009) 106:13641–2. doi: 10.1073/pnas.0906987106
240. Kinsella MT, Monk C. Impact of maternal stress, depression & anxiety on fetal neurobehavioral development. *Clin Obstetr Gynecol*. (2009) 52:425. doi: 10.1097/GRF.0b013e3181b52df1
241. Hoyer D, Nowack S, Bauer S, Tetschke F, Rudolph A, Wallwitz U, et al. Fetal development of complex autonomic control evaluated from multiscale heart rate patterns. *Am J Physiol Regulat Integr Comp Physiol*. (2013) 304:R383–92. doi: 10.1152/ajpregu.00120.2012
242. Van den Bergh BR, Mennes M, Stevens V, van der Meere J, Börger N, Stiers P, et al. ADHD deficit as measured in adolescent boys with a continuous performance task is related to antenatal maternal anxiety. *Pediatr Res*. (2006) 59:78–82. doi: 10.1203/01.pdr.0000191143.75673.52
243. Barker DJ. *In utero* programming of chronic disease. *Clin Sci*. (1998) 95:115–28. doi: 10.1042/cs0950115
244. Värri A, Kallonen A, Helander E, Ledesma A, Pladys P. The Digi-NewB project for preterm infant sepsis risk and maturity analysis. *Finnish J eHealth eWelfare*. (2018) 10:330–3. doi: 10.23996/fjhw.69152
245. Signorini MG, Ferrario M, Cerutti S, Magenes G. Advances in monitoring cardiovascular signals. Contribution of nonlinear signal processing. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Boston, MA: IEEE (2011). p. 6568–71. doi: 10.1109/IEMBS.2011.6091620
246. Magenes G, Signorini MG, Ferrario M, Lunghi F. 2CTG2: A new system for the antepartum analysis of fetal heart rate. In: *11th Mediterranean Conference on Medical and Biomedical Engineering and Computing*. Ljubljana: Springer (2007). p. 781–4. doi: 10.1007/978-3-540-73044-6_203
247. Ferrario M, Signorini MG, Magenes G. New indexes from the Fetal Heart Rate analysis for the identification of severe intra uterine growth restricted fetuses. In: *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. New York, NY: IEEE (2006). p. 1458–61. doi: 10.1109/IEMBS.2006.259274
248. Signorini MG, Magenes G. Reliable nonlinear indices for fetal heart rate variability signal analysis. In: *2014 8th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*. Trento: IEEE (2014). p. 213–4. doi: 10.1109/ESGCO.2014.6847595
249. Lunghi F, Magenes G, Pedrinazzi L, Signorini MG. Detection of fetal distress though a support vector machine based on fetal heart rate parameters. In: *Computers in Cardiology*. Lyon: IEEE (2005). p. 247–50. doi: 10.1109/CIC.2005.1588083
250. Delgado JA, Altuve M, Homsi MN. Haar wavelet transform and principal component analysis for fetal QRS classification from abdominal maternal ECG recordings. In: *2015 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA)*. Bogota: IEEE (2015). p. 1–6. doi: 10.1109/STSIVA.2015.7330451
251. Lu YS, Wei Sy, Liu Xl. Nonlinear FHR baseline estimation using empirical mode decomposition and kohonen neural network. In: *2012 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. Hsinchu: IEEE (2012). p. 368–71. doi: 10.1109/ICoSP.2012.6491896
252. Georgoulas G, Gavrilis D, Tsoulos IG, Stylios C, Bernardes J, Groumpos PP. Novel approach for fetal heart rate classification introducing grammatical evolution. *Biomed Signal Process Control*. (2007) 2:69–79. doi: 10.1016/j.bspc.2007.05.003
253. Inbarani HH, Banu PN, Azar AT. Feature selection using swarm-based relative reduct technique for fetal heart rate. *Neural Comput Appl*. (2014) 25:793–806. doi: 10.1007/s00521-014-1552-x
254. Spilka J, Georgoulas G, Karvelis P, Chudáček V, Stylios CD, Lhotská L. Discriminating normal from "abnormal" pregnancy cases using an automated FHR evaluation method. In: *Hellenic Conference on Artificial Intelligence*. Cham: Springer (2014). p. 521–31. doi: 10.1007/978-3-319-07064-3_45
255. Karvelis P, Spilka J, Georgoulas G, Chudáček V, Stylios CD, Lhotská L. Combining latent class analysis labeling with multiclass approach for fetal heart rate categorization. *Physiol Measure*. (2015) 36:1001. doi: 10.1088/0967-3334/36/5/1001
256. Khandoker AH, Al-Angari HM, Voss A, Schulz S, Kimura Y. Quantification of maternal-fetal cardiac couplings in normal and abnormal pregnancies applying high resolution joint symbolic dynamics. *Math Biosci Eng*. (2020) 17:802–13. doi: 10.3934/mbe.2020042
257. Montero-Nava JE, Pliego-Carrillo AC, Ledesma-Ramirez CI, Pena-Castillo MÁ, Echeverria JC, Pacheco-López G, et al. Analysis of the fetal cardio-electrohysterographic coupling at the third trimester of gestation in healthy women by Bivariate Phase-Rectified Signal Averaging. *PLoS ONE*. (2020) 15:e0236123. doi: 10.1371/journal.pone.0236123
258. Signorini MG, Pini N, Malovini A, Bellazzi R, Magenes G. Integrating machine learning techniques and physiology based heart rate features for antepartum fetal monitoring. *Comput Methods Prog Biomed*. (2020) 185:105015. doi: 10.1016/j.cmpb.2019.105015
259. Lee SK, Park YS, Cha KJ. Recovery of signal loss adopting the residual bootstrap method in fetal heart rate dynamics. *Biomed Eng*. (2019) 64:157–61. doi: 10.1515/bmt-2017-0203

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fped.2021.736834

Detection of Preventable Fetal Distress During Labor From Scanned Cardiotocogram Tracings Using Deep Learning

Martin G. Frasch†, Shadrian B. Strong†, David Nilosek, Joshua Leaverton and Barry S. Schifrin*

Heart Rate AI, Inc., Seattle, WA, United States

Despite broad application during labor and delivery, there remains considerable debate about the value of electronic fetal monitoring (EFM). EFM includes the surveillance of fetal heart rate (FHR) patterns in conjunction with the mother's uterine contractions, providing a wealth of data about fetal behavior and the threat of diminished oxygenation and cerebral perfusion. Adverse outcomes universally associate a fetal injury with the failure to timely respond to FHR pattern information. Historically, the EFM data, stored digitally, are available only as rasterized pdf images for contemporary or historical discussion and examination. In reality, however, they are rarely reviewed systematically or purposefully. Using a unique archive of EFM collected over 50 years of practice in conjunction with adverse outcomes, we present a deep learning framework for training and detection of incipient or past fetal injury. We report 94% accuracy in identifying early, preventable fetal injury intrapartum. This framework is suited for automating an early warning and decision support system for maintaining fetal well-being during the stresses of labor. Ultimately, such a system could enable obstetrical care providers to timely respond during labor and prevent both urgent intervention and adverse outcomes. When adverse outcomes cannot be avoided, they can provide guidance to the early neuroprotective treatment of the newborn.

Keywords: cardiotocography, deep learning-artificial neural network (DL-ANN), fetal brain injury, convolutional neural network (CNN), prevention

OPEN ACCESS

Edited by:
Chan-Wook Park,
Seoul National University, South Korea

Reviewed by:
Mithilesh Lal,
South Tees Hospitals NHS Foundation Trust, United Kingdom
Raghvendra Mall,
St. Jude Children's Research Hospital, United States

****Correspondence:***
Martin G. Frasch
mfrasch@uw.edu;
martin@heartrate.ai

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *05 July 2021*
Accepted: *21 October 2021*
Published: *03 December 2021*

Citation:
Frasch MG, Strong SB, Nilosek D, Leaverton J and Schifrin BS (2021) Detection of Preventable Fetal Distress During Labor From Scanned Cardiotocogram Tracings Using Deep Learning. Front. Pediatr. 9:736834. doi: 10.3389/fped.2021.736834

INTRODUCTION

In the United States, there are approximately four million births per year (1). Over 85% of them are accompanied by electronic fetal monitoring (EFM) in labor with the objective of safeguarding fetal/neonatal well-being. This surveillance of the FHR pattern (rhythm) in conjunction with the mother's uterine contractions provides a wealth of data about fetal behavior and the threat of diminished oxygenation and cerebral perfusion. Fifty years after its introduction, however, fetal monitoring continues to inspire debate about its value and especially its role in the increasing cesarean section rate as well as being a "litogen"—a stimulus to allegations of medical malpractice (2–10). Reviews of adverse labor outcomes in numerous countries universally associate adverse fetal outcomes with the failure to timely respond to the FHR pattern information [(11, 12); Inquiries, personal communication]. Indeed, various sources affirm that misinterpretation of EFM (or the uncertainty with patterns) has

contributed to the significantly increased use of cesarean delivery from 5% in the 1970s to >30% today (13, 14), leading to increased expenditures, incurring costs in the United States (13, 14) of over $1 billion per year per 5% of additional cesarean deliveries (15). Obstetrical liability costs the country ∼$40 billion per year, of which 70% is accounted for by uncertainty about EFM interpretation and related brain injury (14, 15).

Earlier and more precise recognition of the precursors of fetal compromise and the institution of corrective/preventative initiatives during labor are urgently needed. Only rarely should urgent delivery be required (16). Additional benefits include better maternal and child outcomes thanks to the avoidance of early intervention, lower cesarean delivery rate, and immediate neonatal monitoring of heart rate pattern, i.e., having the baby continuously monitored for at least 15 min after delivery. Here, babies seen to be at risk can be evaluated and more aggressively treated earlier than currently undertaken.

Historically, the EFM data, stored digitally, are available only as rasterized pdf images for contemporary or historical discussion and examination (**Figure 1**). In reality, however, they are rarely reviewed systematically or purposefully. In the case of a medical–legal review, it is the paper copy of the tracing, exclusively, that is likely available and consulted.

We propose a deep learning (DL)-based approach to this challenge. It is based on a unique archive which collected over four decades of EFM tracings of babies with known, adverse outcomes. This archive provides many unique examples of the broad range of healthy, threatened, and injured fetuses along with their long-term follow-up. Consequently, this archive is ideal for automating an early warning (preventive guidance) system for maintaining fetal well-being during the stresses of labor and delivery that could ultimately enable a health care provider to timely and conservatively respond during labor to prevent urgent interventions and adverse outcomes. When adverse outcomes cannot be avoided, they guide the early neuroprotective treatment of the newborn. This system utilizes a unique classification of heart rate and contraction patterns (details in section Methods), including specific identifiable indicators ("point A" and "point B") of the need for attention by the provider (16–19).

METHODS

Data

For this pilot study, a convenience sample of 36 tracings was selected. All tracings were derived from singleton pregnancies at term undergoing a trial of labor with a fetal monitor in place as previously described (18). Each tracing was considered normal at the onset of monitoring—an important distinction. The majority of features were derived from conventional guidelines (ACOG) including baseline rate, variability, accelerations, and decelerations. For this study, however, certain operational definitions of heart rate patterns (**Table 1**) and uterine contractions (**Table 2**) were modified by the subject matter

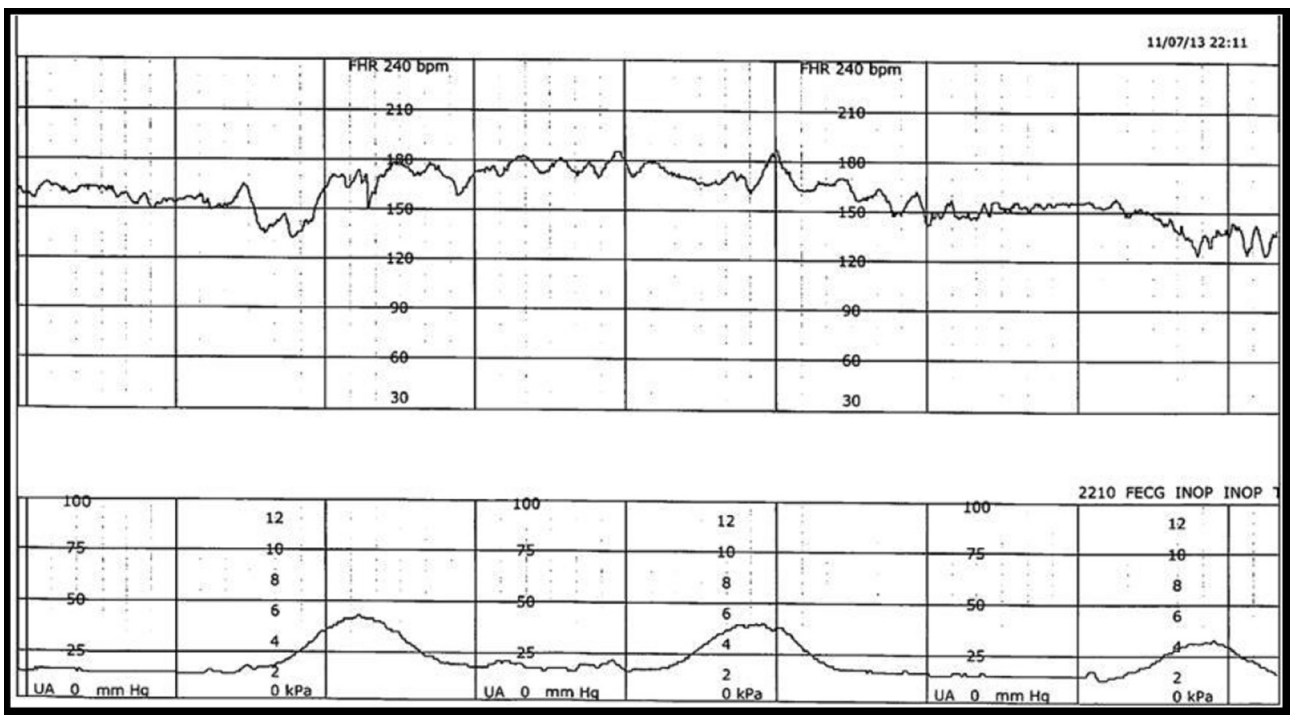

FIGURE 1 | Example of FHR (top) and uterine contraction (bottom) during labor, captured simultaneously and stored electronically in a digital format but available only as a rasterized pdf document.

TABLE 1 | Definitions of EFM patterns.

Basal heart rate	The baseline FHR established at the beginning of labor with fetus quiescent
Tachycardia	Absolute—sustained (>10 min) baseline heart rate above 155 bpm
	Relative—sustained (>10 min) baseline heart rate >15 above basal rate
Bradycardia	Absolute—sustained (>10 min) baseline heart rate below 110 bpm
	Relative—sustained (>10 min) baseline heart rate >15 bpm below the basal rate
Deceleration recovery	The response of the fetus to a deceleration
	Categories of recovery:
Normal response	Prompt return to the previously normal baseline rate and variability
Adverse response	Applies to the recovery of the deceleration but may persist as a feature of the subsequent baseline heart rate
Overshoot	An acceleration of the FHR immediately following a deceleration with a duration proportional to the amplitude of the preceding deceleration. Usually associated with alterations in baseline rate and variability
Delayed return	A "slow return" to the baseline—likely a sustained elevation of fetal blood pressure in anticipation of recovery
Peaked return	An abrupt peak at the end of a deceleration followed by a late deceleration. An ominous commentary usually leading to fetal death
Decreased/absent variability	Persistent diminution in baseline variability <6 bpm
Increased variability	Persistent or transient elevation of variability >25 bpm
Sinusoidal pattern	Visually apparent, smooth, sine wave-like undulating pattern in FHR baseline with a cycle frequency of 3–5 per min. Occurs in the absence of a normal CTG pattern nearby. May be brief or persistent
Checkmark pattern	A unique pattern seen in neurologically compromised/asphyxiated fetuses suggesting repetitive "checkmarks" () of varying duration—frequently elicited by a preceding deceleration
Sawtooth pattern	Rapid, high frequency (20+ cpm), low amplitude (<15 bpm), peaked oscillations in the heart rate that generally increase in frequency and decrease in amplitude over time
Conversion	A CTG pattern in which there is a dramatic change in rate, variability, and pattern of deceleration within 1–2 contractions—suggests fetal ischemic injury

TABLE 2 | Definition of excessive uterine activity.

Contraction parameter	Average	Excessive
Frequency	2–4.5 UC/10 min	>5/10 min (×2)
Intensity	25–75 mmHg	Not defined
Duration	60–90 s	>90 s
Resting tone	12–20 mmHg	>20 mmHg
Interval between peaks	2–4 min	<120 s
Rest time*	50–75%	<50%
Montevideo units	Not used	

**Rest time—interval when contractions and pushing are absent.*
UC, uterine contractions; mmHg, millimeters of mercury.

expert (**Figure 2**). These included the basal rate, the use of relative bradycardia and tachycardia, and the pattern of recovery of the deceleration.

Identification of EFM Features

Tracing is defined at the outset of monitoring as normal or abnormal.

A normal tracing is characterized by a stable baseline heart rate between 110 and 155 bpm, with moderate variability and absent decelerations. An abnormal tracing is characterized by at least one of the following features:

- baseline heart rate: <100, >155, arrhythmia, unstable/indeterminate;
- baseline variability: absent, decreased (<6 bpm), increased (>25 bpm); and
- decelerations: late, variable, undefined.

Thus, for decelerations with normal recovery, no immediate action is required. They return promptly to the previously normal baseline variability (5–15 bpm peak to trough and chaotic, pseudorandom) and heart rate (usually 110–155 bpm and stable); each fetus has an individually unique baseline (basal rate). A "normal" deceleration returns to baseline without changing trajectory, and upon reaching the previous baseline rate remains there. These features pertain regardless of amplitude, duration, and timing of the deceleration and signify the comfortable compensation for the alteration in blood flow represented by the deceleration.

Point A

Point A denotes the time when the recovery of the deceleration is no longer "normal" and those additional compensatory activities are invoked by the fetus to maintain homeostasis. The detection of Point A signifies that increased attention and conservative measures are needed in an attempt to restore homeostasis to the previously normal tracing. These features include the following:

A. Delayed return to baseline: includes a change in the trajectory of the recovery such that the return to baseline is delayed beyond the end of the contraction.
B. Period of increased variability: peak to trough >20 bpm, frequency 5–10 cycles per min. Duration is also influenced by the appearance of a subsequent contraction during which time the pattern disappears—taken over by the deceleration.
C. Overshoot: an acceleration following the upslope of the return of the deceleration lasting 15 s or more prior to the return to the baseline.
D. Transient (usually at least 1 min) return to a higher baseline by at least 15 bpm, duration affected by next contraction, compared to the previously stable baseline.

E. Transient (at least 1 min) return to a lower baseline by at least 10 bpm compared to the previously normal baseline.
F. Excessive uterine activity—irrespective of changes in FHR pattern.

The detection of Point A alerts the health care provider to the need for at least conservative intervention in regard to the maternal condition, the frequency of contractions, or expulsive efforts during the second stage of labor. Point A is identified sooner if an excessive uterine activity is present.

Point B

Point B represents the attempt to define significant fetal compromise or injury, irrespective of the perceived amount of acidosis (pH) in the fetus. No clinical circumstances are used in the definition of Point B. Point B was identified by the subject expert (BSS) *via* a custom-created digital interface (AWS) allowing us to feed the annotations directly into the DL model.

These features include the following:

- sustained return to a baseline with diminished/absent baseline variability, usually accompanied by a rise in the baseline heart rate; and
- sustained change in baseline rate and variability with adverse features (**Table 1**) occurring within 5 to 10 min of a previously normal rate and variability—usually with recurrent variable decelerations.

DL Pipeline

We present a method for automated extraction of features in FHR and uterine contractions (UCs), which are outlined in the above section.

Briefly, the method includes acquiring a set of non-digitized charts, digitally assigning markers to predetermined features in the charts, supplying the assigned marker-feature sets to a supervised model, statistically iterating over the assigned sets, automatically assessing model performance, and applying the model to new sets of charts to extract non-assigned predetermined features.

The method for automated chart processing includes analyzing time-series sets of non-digitized charts of FHR and/or concurrent maternal UC to digitally associate markers with fetal signatures, using the associated groups for supervised training of an artificial intelligence model, determining accuracy and precision of the model, and applying the trained model to automatically process new time-series sets of one or more charts of FHR and concurrent maternal UC, having one or more unassociated fetal signatures.

To achieve this goal, we treat scanned EFM recordings as non-vectorized images, similar to digital photographs, and apply supervised machine learning to extract and process features to train an artificial intelligence model. An image is supplied to a convolution neural network (CNN) model (20). The image is represented as one or more numeric arrays of pixel values with varying signal counts associated with the pixel content. The pixel content is dictated by the amount of red, green, blue, or other spectral bands that the pixel may receive and is an integer number in one or more dimensions. The CNN is represented as a set of algorithmic layers into which the numeric pixel data are sent. It consists of a series of convolutional layers, non-linear layers, pooling layers, and fully connected layers. Each such layer may be considered an individual set of equations, where the output of one equation becomes the input to another. The CNN eliminates the need for manual feature extraction, as the features are acquired through the passing of the pixel data to one or more other layers, and correlations are extracted and weighted as a consequence of the layer transitions.

We implement a single-shot detector (SSD) algorithm to achieve this goal (21). It utilizes a standard CNN network (e.g., VGG-16) with an additional set of convolution layers to identify discrete locations of one or more features in one or more images (22). The SSD codebase is available here: https://github.com/zhreshold/mxnet-ssd.

Through a single pass in the CNN, the weighted correlations meant to describe the relevant features are tested against ground truth data (validation data), separate from training data. The goal of this statistically iterative operation is to minimize a loss function between the predicted correlations and the truth values through adaptively updating the weights of the predicted function. The process of adjusting the weights continues until a minimum statistical loss is obtained.

The output model and weights are then used for inference against the withheld (unseen) dataset to extract similar relevant information.

Sample Selection and Processing

Briefly, in this study, we implemented a conventional random 80/20 train/test split. This corresponded to 26.4 h of training on EMF image information and 6.6 h used for testing (validation).

That is, the EFM images were flipped/translated, and the noise was added to represent more of the variability observed in the original pdfs.

The samples used in the analysis were 36 unique medical case reports in the form of a static pdf. The pdfs were split and converted into individual PNG images, one PNG per page in the pdf. As each pdf report consisted of a different number of unique pages, the number of images per page varied. In the end, there were 252 image pages across the 36 individual medical cases. The images were further cropped automatically to contain just the graphical data component of the page, removing headers, footers, and extraneous text. This was then split into 80/20 train/test datasets, resulting in 202 training image graphs and 50 test image graphs. Of the 50 test images, 25 were held back for separate validation. The images consolidated in these training and testing datasets were similar in quality (bold graphs with discernible FHR features). It should be noted that many additional pdf reports contained a varying degree of quality based on the photocopied/scanned/faxed nature of the captured data. This presented a significant challenge to create a robust training dataset with representative features. With 202/50 train/test data, significant augmentation was required. The images were flipped in horizontal space, as that preserved the domain of the information. A vertical flip would manifest in features unrepresentative of the FHR signatures. A further augmentation was required to reduce or sharpen the resolution

TABLE 3 | Cohort characteristics.

	Age, years	EGA, weeks	BMI	BWT, g	Apgar 1	Apgar 5
Median	26.5	39.8	31.0	3,325	2	6
25th	21.0	39.1	26.7	3,070.0	1.0	4.0
75th	30.3	40.4	35.9	3,601.8	4.0	7.0
	Temporal characteristics of labor (h:min)					
	1st	2nd	Labor	Point A	Point A to Delivery	Point B	Point B to Delivery
Median	14:24	3:35	23:10	13:38	4:01	10:30	0:43
25th	1:30	2:49	16:33	5:44	1:52	4:20	0:23
75th	13:30	5:20	10:12	20:43	5:18	18:20	1:37

of the images, to better capture the variability of the pdf graphs. In the end, 808 images were used in training and 200 in testing. These data are still rather shallow for DL, as the feature space and EFM signatures possible are vast. In future work, further data must be included in order to fully represent the feature space of the variabilities of both the documents and the EFM signatures. As an SSD algorithm was leveraged to isolate the EFM events, the images were manually annotated with standard data labeling practices, and output into the Pascal VOC XML format (http://host.robots.ox.ac.uk/pascal/VOC/).

RESULTS

The demographics and clinical characteristics are summarized in **Table 3**. There were 11 outcomes with a pH <7.00. **Table 3** also denotes the duration and timing of the first and second stages of pushing, Point A and Point B.

There were numerous points in the dataset that were abnormal but did not trigger Point A. Isolated but persistent changes in baseline rate, baseline variability, and excessive uterine activity are commonplace and do make the tracing abnormal without evoking Point A. Eventually, Point A was reached in all instances in this dataset. As such, from a machine learning perspective, this is a balanced dataset. This is also implied by the column "Point A to delivery" (**Table 3**).

As a step toward developing this proactive fetal surveillance system, we have created an artificial intelligence model using a basic SSD DL approach to retrospectively identify critical features in the EFM data (cardiotocography) from the rasterized pdf directly (**Figure 3**). This model creates a classification of the pattern and identifies critical features of the tracing that indicate critical and timely points of either conservative or operative intervention, "Point A" and "Point B." Here, in the initial implementation, we focused on predicting "Point A."

This novel application of using pdf rasterized plots as an image detection DL problem facilitates (1) quick and efficient deployment against a large record of data without chart digitizing and (2) packaging and deployment as a lightweight or MobileNet (23) application useful for immediate integration with a mobile device, post event.

The model achieved an accuracy of 93.6% in identifying Point A (i.e., detecting accurately the entire test set of features comprising Point A) against a small dataset with limited variability in features.

The average intersection over union (IoU) for the 25 validation images was 0.67, indicating a 67% overlap in the area with the true annotated feature. Annotated features are described in **Table 1**. This was averaged over 47 EFM bounding box features (true features and negative features) in the 25 images. Of the 47 features, the precision and recall were 87 and 82.5%.

DISCUSSION

Our primary goal was the early identification of abnormal tracings at the outset (considered Point A) and the early detection of isolated adverse features (abnormal) whose coalescence (Point A) demands intervention at a time when correction is likely. We successfully implemented automated identification of Point A, indicating threatened fetal decompensation of the highest relevance for real-time clinical implementation of such an algorithm. The SSD approach we deployed uses baseline data to identify Point A. In other words, the expert diagnosis of Point A on which the model was trained takes the baseline into account and seeks to identify the patterns comprising Point A in relation to the baseline.

In response to Point A, conservative rehabilitative measures include the following:

- diminishing the frequency of uterine contractions;
- diminishing/ceasing pushing during the second stage of labor;
- decreasing infusion of oxytocin; and
- assessing the feasibility of safe vaginal delivery.

However, the suggested measures cannot be ranked in relation to the probability score of Point A that our model provides as their sequence is primarily responsive to the feature(s) that prompted the response. If the problem involves excessive uterine activity, the care provider is directed to diminish uterine activity. If the response reveals late decelerations, the care provider is directed to modify the patient's position, providing supplemental oxygen, assisting with maternal blood pressure, etc.

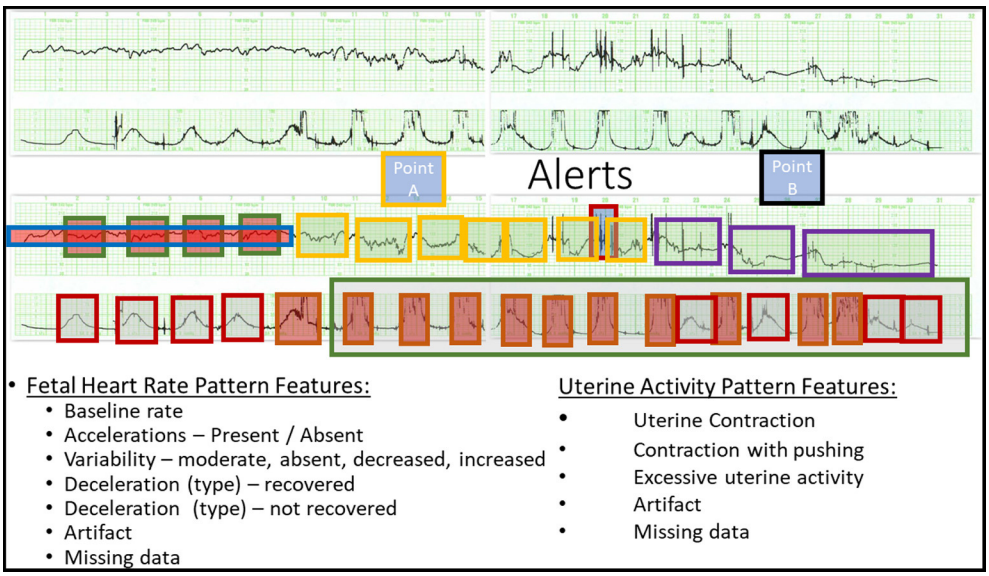

FIGURE 2 | Definition of Point A and Point B. (Top) A representative raw CTG tracing. (Bottom) The annotated CTG tracing deriving Point A and Point B. See **Tables 1, 2** for details.

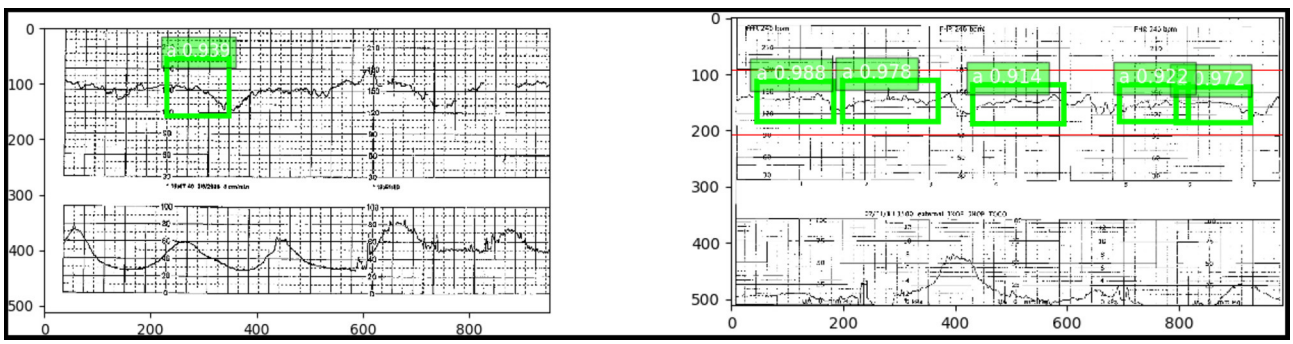

FIGURE 3 | (Left) "Point A" identified with a 93.6% accuracy using an SSD trained on the pdf chart images. (Right) Numerous occurrences of "Point A" with high confidence in green. Red indicates the true "Point A" duration.

Based on our observations of ∼5,000 cases with brain injury as the birth outcome, 20% of normal patients reach Point A. About 25% of these revert to normal. Point B is reached in about 0.5% of the population and in about 30–40% of our observations on brain-injured babies with subsequent handicaps. We leave it to future work to implement the prediction of Point B, as this will require training on larger datasets. These points, together with other key signs in the FHR, can be displayed for the obstetrical care provider as part of an early alert and decision support system. Consequently, the visual signature for training the SSD is extracted similarly to the method utilized by the physician. The time-series nature of the FHR may be exploited with an additional application of a long short-term memory (LSTM) (24) model for consistent identification and tracking as a function of event duration. However, to date, only the SSD has been deployed.

It is important to emphasize that the training of the model was not based on the detection of acidosis or even low Apgar score, but whether or not conservative intervention based on the cardiotocographic pattern (Point A) was deemed necessary and whether criteria were met for the presumptive diagnosis of fetal neurological injury (Point B) as described previously (17). There was no attempt to correlate the outcome results with either pH or Apgar score of the newborn.

It may be seen as a limitation of the study that we did not seek correlations with fetal acidosis, Apgar scores, need for resuscitation, or NICU admission for HIE. However, the objective of our study was to use DL to prevent urgent intervention ("rescue") by identifying the point in the previously normal tracing before fetal acidosis has developed and where conservative measures can be expected to restore the tracing to normal. We see no benefit in employing an artificial intelligence system to detect acidosis and, simultaneously, the need to rescue

the fetus that may have already become injured. The system is designed to work with fetuses with initially normal tracings as no real benefit can be calculated from an algorithm that begins with an abnormal tracing where the options for prevention are limited and early delivery is likely (25).

Another limitation is that here we deliver a proof of concept only, using a convenience sample of 36 tracings in singleton pregnancies only. We leave a validation on the larger dataset and in multiple pregnancies, preterm deliveries, or IUGR fetuses for future work.

The approach to presenting and interpreting existing clinical data and annotating the EFM record during labor can dramatically reduce the need for urgent deliveries and significantly improve the outcomes of babies and mothers in labor and for the neonate in the nursery. Improved outcomes, less urgency, fewer rescues, and better documentation could be a game-changer for the care of pregnant women and children and the defense of allegations of obstetrical malpractice.

In future work, to boost the present performance results, alternatively or additionally, RCNN, LSTM, RNN, support vector machine, random forest, instance segmentation, image classification techniques, and/or other DL algorithms and/or other machine learning techniques can be applied.

The new EFM data can be supplied to the trained model in a format different from the format of the original training/testing images. For example, the EFM data can be supplied in the format of digitized charts, tabularly represented data, a signal received from one or more devices, etc. In other words, once the model has been trained, it can be configured to work on similar features provided in the same and/or other data formats, including live data. Such an approach allows the model to identify one or more features of interest and also the location of those features in the chart(s). This location can be correlated with a time and/or other dependent variables within the chart and/or a set of charts.

These features of our approach make it attractive to electronic medical record and physiological monitoring applications well beyond EFM.

We have shown the feasibility of a DL approach to scan and detect the ability of the fetus to handle the trial of labor using standard FHR and uterine activity chart tracings presented to artificial intelligence in the form of images, the format in which the majority of such tracings are still stored and presented to the experts for the determination of the need for intervention and the timing of the fetal injury. Our DL approach detects these factors with over 90% accuracy (compared to expert scoring).

DATA AVAILABILITY STATEMENT

The SSD codebase is available here: https://github.com/zhreshold/mxnet-ssd.

AUTHOR CONTRIBUTIONS

MGF, SS, DN, and BS conceived the manuscript. MGF wrote the initial draft. BS, SS, and MGF conducted the analyses. MGF, SS, DN, JL, and BS revised and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

1. CDC. *Births and Natality*. (2021). Available online at: https://www.cdc.gov/nchs/fastats/births.htm (accessed September 27, 2021).
2. Graham EM, Adami RR, McKenney SL, Jennings JM, Burd I, Witter FR. Diagnostic accuracy of fetal heart rate monitoring in the identification of neonatal encephalopathy. *Obstet Gynecol*. (2014) 124:507–13. doi: 10.1097/AOG.0000000000000424
3. Alfirevic Z, Devane D, Gyte GM, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev*. (2017) 2:CD006066. doi: 10.1002/14651858.CD006066.pub3
4. Hirsch E. Electronic fetal monitoring to prevent fetal brain injury: a ubiquitous yet flawed tool. *JAMA*. (2019) 322:611–612. doi: 10.1001/jama.2019.8918
5. Sartwelle TP. Defending a neurologic birth injury. Asphyxia neonatorum redux. *J Leg Med*. (2009) 30:181–247. doi: 10.1080/01947640902936522
6. Sartwelle TP, Johnston JC. Cerebral palsy litigation: change course or abandon ship. *J Child Neurol*. (2015) 30:828–41. doi: 10.1177/0883073814543306
7. Afors K, Chandraharan E. Use of continuous electronic fetal monitoring in a preterm fetus: clinical dilemmas and recommendations for practice. *J Pregnancy*. (2011) 2011:848794. doi: 10.1155/2011/848794
8. Pinas A, Chandraharan E. Continuous cardiotocography during labour: analysis, classification and management. *Best Pract Res Clin Obstet Gynaecol*. (2016) 30:33–47. doi: 10.1016/j.bpobgyn.2015.03.022
9. Frasch MG, Boylan GB, Wu HT, Devane D. Commentary: computerized interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Front Physiol*. (2017) 8:721. doi: 10.3389/fphys.2017.00721
10. Frasch MG. Saving the brain one heartbeat at a time. *J Physiol*. (2018) 596:5503–4 doi: 10.1113/JP275776
11. Resolution NHS. *The Early Notification Scheme Progress Report: Collaboration and Improved Experience for Families*. (2019). Available online at: https://resolution.nhs.uk/wp-content/uploads/2019/09/NHS-Resolution-Early-Notification-report.pdf (accessed September 2019).
12. Chandraharan E, Arulkumaran S. Prevention of birth asphyxia: responding appropriately to cardiotocograph (CTG) traces. *Best Pract Res Clin Obstet Gynaecol*. (2007) 21:609–624. doi: 10.1016/j.bpobgyn.2007.02.008
13. Betrán AP, Ye J, Moller AB, Zhang J, Gülmezoglu AM, Torloni MR. The increasing trend in caesarean section rates: global, regional and national estimates: 1990-2014. *PLoS ONE*. (2016) 11:e0148343. doi: 10.1371/journal.pone.0148343
14. Institute of Medicine (US) Committee to Study. Medical professional liability and the delivery of obstetrical car. In: Rostow VP, Bulger RJ, editors. *Is the Rising Rage of Cesarean Sections a Result of More Defensive Medicine?* National Academies Press (US) (1989). Available online at: https://www.ncbi.nlm.nih.gov/books/NBK218656/ (accessed November 27, 2019).
15. Institute of Medicine (US) Committee to Study. Medical professional liability and the delivery of obstetrical care. In: Rostow VP, Bulger RJ, editors. *Medical Professional Liability and the Delivery of Obstetrical Care: Volume II: An Interdisciplinary Review*. Washington, DC: National Academies Press (US) (2014). p. 78–96.
16. Eden RD, Evans MI, Britt DW, Evans SM, Schifrin BS. Safely lowering the emergency Cesarean and operative vaginal delivery rates using the fetal reserve index. *J Matern Fetal Neonatal Med*. (2018) 1–7:1473–9. doi: 10.1080/14767058.2018.1519799
17. Schifrin BS, Ater S. Fetal hypoxic and ischemic injuries. *Curr Opin Obstet Gynecol*. (2006) 18:112–22. doi: 10.1097/01.gco.0000192984.15095.7c

18. Eden RD, Evans MI, Evans SM, Schifrin BS. The "Fetal Reserve Index": re-engineering the interpretation and responses to fetal heart rate patterns. *Fetal Diagn Ther*. (2018) 43:90–104. doi: 10.1159/000475927
19. Eden RD, Evans MI, Evans SM, Schifrin BS. Reengineering electronic fetal monitoring interpretation: using the fetal reserve index to anticipate the need for emergent operative delivery. *Reprod Sci*. (2018) 25:487–97. doi: 10.1177/1933719117737849
20. Petrozziello A, Jordanov I, Aris Papageorghiou T, Christopher Redman WG, Georgieva A. Deep learning for continuous electronic fetal monitoring in labor. *Annu Int Conf Proc IEEE Eng Med Biol Soc*. (2018) 2018:5866–9. doi: 10.1109/EMBC.2018.8513625
21. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, et al. SSD: single shot MultiBox detector. In: *Computer Vision – ECCV 2016*. Amsterdam: Springer International Publishing. (2016). p. 21–37.
22. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv [csCV]* (2014). Available online at: http://arxiv.org/abs/1409.1556
23. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. *arXiv [csCV]* (2017). Available online at: http://arxiv.org/abs/1704.04861
24. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. (1997) 9:1735–80. doi: 10.1162/neco.1997.9.8.1735
25. Jonsson M, Ågren J, Nordén-Lindeberg S, Ohlin A, Hanson U. Neonatal encephalopathy and the association to asphyxia in labor. *Am J Obstet Gynecol*. (2014) 211:667.e1–8. doi: 10.1016/j.ajog.2014.06.027

Conflict of Interest: MGF, SS, DN, JL, and BS are co-founders of Heart Rate AI Inc. The research has been conducted as part of product development by Heart Rate AI Inc. and all work has been made open-source.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pediatrics

ORIGINAL RESEARCH
published: 16 March 2022
doi: 10.3389/fped.2022.784439

Cardiotocography and Clinical Risk Factors in Early Term Labor: A Retrospective Cohort Study Using Computerized Analysis With Oxford System

Aimée A. K. Lovers[1], *Austin Ugwumadu*[2] *and Antoniya Georgieva*[1]*

[1] *Nuffield Department of Women's and Reproductive Health, Big Data Institute, University of Oxford, Oxford, United Kingdom,*
[2] *Department of Obstetrics and Gynaecology, St George's, University of London, London, United Kingdom*

Objective: The role of cardiotocography (CTG) in fetal risk assessment around the beginning of term labor is controversial. We used routinely collected clinical data in a large tertiary hospital to investigate whether infants with "severe compromise" at birth exhibited fetal heart rate abnormalities in their first-hour CTGs and/or other clinical risks, recorded as per routine care.

Materials and Methods: Retrospective data from 27,927 parturitions (single UK tertiary site, 2001–2010) were analyzed. Cases were included if the pregnancy was singleton, ≥36 weeks' gestation, cephalic presentation, and if they had routine intrapartum CTG as per clinical care. Cases with congenital abnormalities, planned cesarean section (CS), or CS for reasons other than "presumed fetal compromise" were excluded. We analyzed first-hour intrapartum CTG recordings, using intrapartum Oxford System (OxSys) computer-based algorithms. To reflect the effect of routine clinical care, the data was stratified into three exclusive groups: infants delivered by CS for "presumed fetal compromise" within 2 h of starting the CTG (*Emergency CS*, $n = 113$); between 2 and 5 h of starting the CTG (*Urgent CS*, $n = 203$); and the rest of deliveries (*Others*, $n = 27{,}611$). First-hour CTG and clinical characteristics were compared between the groups, subdivided to those with and without severe compromise: a composite outcome of stillbirth, neonatal death, neonatal seizures, encephalopathy, resuscitation followed by ≥48 h in neonatal intensive care unit. Two-sample *t*-test, X^2 test, and Fisher's exact test were used for analysis.

Results: Compared to babies without severe compromise, those with compromise had significantly higher proportion of cases with baseline fetal heart rate ≥150 bpm; non-reactive trace; reduced long-term and short-term variability; decelerative capacity; and no accelerations in the first-hour CTG across all groups. Prolonged decelerations(≥3 min) were also more common. Thick meconium and small for gestational age were consistently more common in compromised infants across all groups. There was more often thick meconium, maternal fever ≥38 C, sentinel events, and other clinical risk factors in the *Emergency CS* and *Urgent CS* compared to the *Others* group.

OPEN ACCESS

Edited by:
Robert Galinsky,
Hudson Institute of Medical Research, Australia

Reviewed by:
Serafina Perrone,
University of Parma, Italy
João Francisco Bernardes,
University of Porto, Portugal

***Correspondence:**
Antoniya Georgieva
antoniya.georgieva@wrh.ox.ac.uk

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *27 September 2021*
Accepted: *14 January 2022*
Published: *16 March 2022*

Citation:
Lovers AAK, Ugwumadu A and Georgieva A (2022) Cardiotocography and Clinical Risk Factors in Early Term Labor: A Retrospective Cohort Study Using Computerized Analysis With Oxford System. Front. Pediatr. 10:784439. doi: 10.3389/fped.2022.784439

Conclusion: A proportion of infants born with severe compromise had significantly different first-hour CTG features and clinical risk factors.

Keywords: cardiotocography, CTG, electronic fetal monitoring, hypoxic-ischaemic encephalopathy, HIE, big data

INTRODUCTION

Intrapartum hypoxic ischaemic injury is a major contributor to stillbirths, neonatal encephalopathy and mortality, and long-term neurodevelopmental sequelae worldwide. Its prevention is an important public health priority (1–3). Over half of term stillbirths and brain injuries around birth are potentially avoidable by better obstetric care (3–6). Efforts to improve the identification of fetuses at-risk of intrapartum compromise are undermined by the imprecision of currently available tools including the conventional and non-computerized cardiotocograph (CTG) (6–8). The CTG is widely used for assessment of fetal wellbeing and has been the mainstay of fetal monitoring for over 50 years (9, 10). However, CTG interpretation is subjective and relies on the use of guidelines derived from imprecise and static patterns of the fetal heart rate (FHR). Unsurprisingly, delays in intervention and avoidable harm may occur, leading to litigation (10–14). The CTG is also associated with increased operative delivery and healthcare costs (11–13). Therefore, there is increasing interest in more objective and reliable methods for CTG interpretation. This includes computer-based analysis based on a very large routinely collected dataset of CTG and maternity data, such as the Oxford System (OxSys) that is currently under development (15). Moreover, automated CTG evaluation with OxSys enables the analysis of large routinely collected sets of CTG traces, providing evidence for the association of different CTG characteristics, labor management, and labor outcomes, which was the focus of this work.

In current practice, the risk of adverse labor outcomes is based on the assessment of antenatal factors such as abnormal fetal growth, antepartum hemorrhage, prolonged rupture of fetal membranes, meconium staining of the amniotic fluid, or abnormal CTG pattern in the initial assessment of women around labor onset (16, 17). However, the pathophysiology of fetal compromise is complex, multifactorial, and dynamic, and the fetal ability to adapt to oxygen deprivation is significantly challenged by labor and underlying conditions such as placental insufficiency, maternal diabetes, and intrauterine infection, which can weaken the fetal adaptation to oxygen deprivation (18, 19). Furthermore, the role of the CTG in the initial assessment of women around the onset of labor is controversial, because different endpoints are used. Previous randomized controlled trials (RCTs) investigated whether an "admission test" *predicted* neonatal outcome at delivery in the so-called low-risk pregnancy. The primary outcome varied in these different RCTs, including cord acidosis at birth, need for continuous CTG, overall perinatal mortality and morbidity, and operative delivery rates (17, 20, 21). Nonetheless, the original "admission test" aimed to identify the fetus with pre-existing compromise in order to institute obstetric intervention (22). Parts et al. (20) subsequently investigated this original aim in a large retrospective observational study of routine admission CTG in both low- and high-risk pregnancies. The authors investigated "all women who underwent emergency cesarean section due to suspected fetal distress" within 1 h of admission.

Following the objective of the original "admission test" and the selected endpoints by Parts et al., we consider that a more appropriate endpoint for studies of the role of CTG assessment at admission is the early identification of fetuses who are already compromised or vulnerable (e.g., fetal growth restriction, infection, and feto-maternal hemorrhage) and who would benefit from early cesarean delivery in order to avoid further compromise during labor. Therefore, this study aims to evaluate the association between severely compromised infants at birth and early warning signs in their first-hour CTG (recorded before or in first-stage labor as per clinical care) and/or clinical risks, stratifying for those with emergency cesarean section performed shortly after admission due to presumed fetal compromise. We used computerized and classic CTG features, using the OxSys computer-based algorithms, to study this large cohort of infants.

MATERIALS AND METHODS

This was a retrospective cohort study of infants delivered at the John Radcliffe Hospital in Oxford, UK, using a clinical data collection system between January 2001 and December 2010. The study received ethical approval from the Newcastle & North Tyneside 1 Research Ethics Committee, Reference 11/NE0044 (data before 2008), and from the South Central Ethics Committee, Reference 13/SC/0153 (for data beyond 2008). Informed consent by the participants was not required.

Data

During the 10-year study period, there were 32,743 singleton deliveries (excluding elective cesarean sections) with gestational age $\geq$ 36 weeks of cephalic presentation and normally formed fetuses with intrapartum CTG recorded as per standard care at the John Radcliffe Hospital, Oxford (**Figure 1**). The selected gestational age of 36 completed weeks is in line with the threshold used in our previous work and in large randomized controlled trials on intrapartum CTG (23–25). We only included cases with at least 20 min of CTG recordings of acceptable signal quality for analysis, taken only at maternity admission or delivery units. The national guidance at the time and still

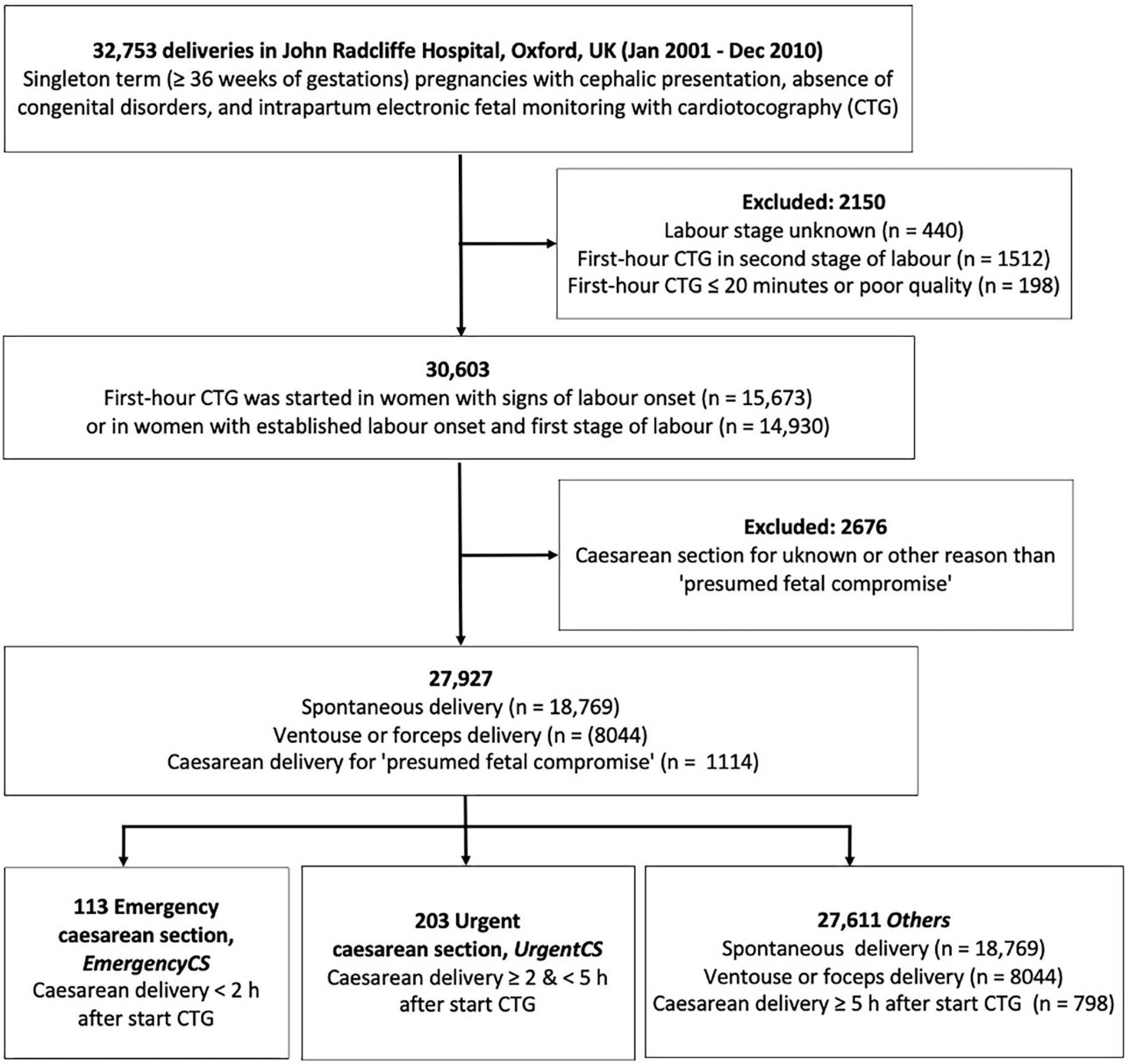

FIGURE 1 | Data flowchart.

is intermittent auscultation for the intrapartum surveillance of low-risk pregnancies, defined as the absence of any antenatal or intrapartum, fetal, maternal, placental, medical, or obstetric complications or risk factors. **Figure 1** shows the study data flow chart. We excluded cases where the CTG recording commenced in the second stage of labor and those with cesarean delivery for indication other than "presumed fetal compromise." A total of 27,927 deliveries were eligible for inclusion and analysis.

Methods

All data were routinely collected as part of standard clinical care and analyzed retrospectively. Outcome groups were based on the mode and timing of cesarean delivery for presumed fetal compromise (**Figure 1**). Infants delivered by cesarean section within 2 h of CTG monitoring (emergency cesarean delivery, *Emergency CS*) were compared with their counterparts delivered by cesarean section between 2 and 5 h (urgent cesarean delivery, *Urgent CS*). The remaining births formed the control group (*Others*).

Initial assessment was defined as the assessment of fetal wellbeing and clinical risk factors in women presenting to the unit with symptoms and signs of labor.

Sentinel events included placental abruption, uterine rupture, and cord prolapse. Birth trauma included intracranial laceration

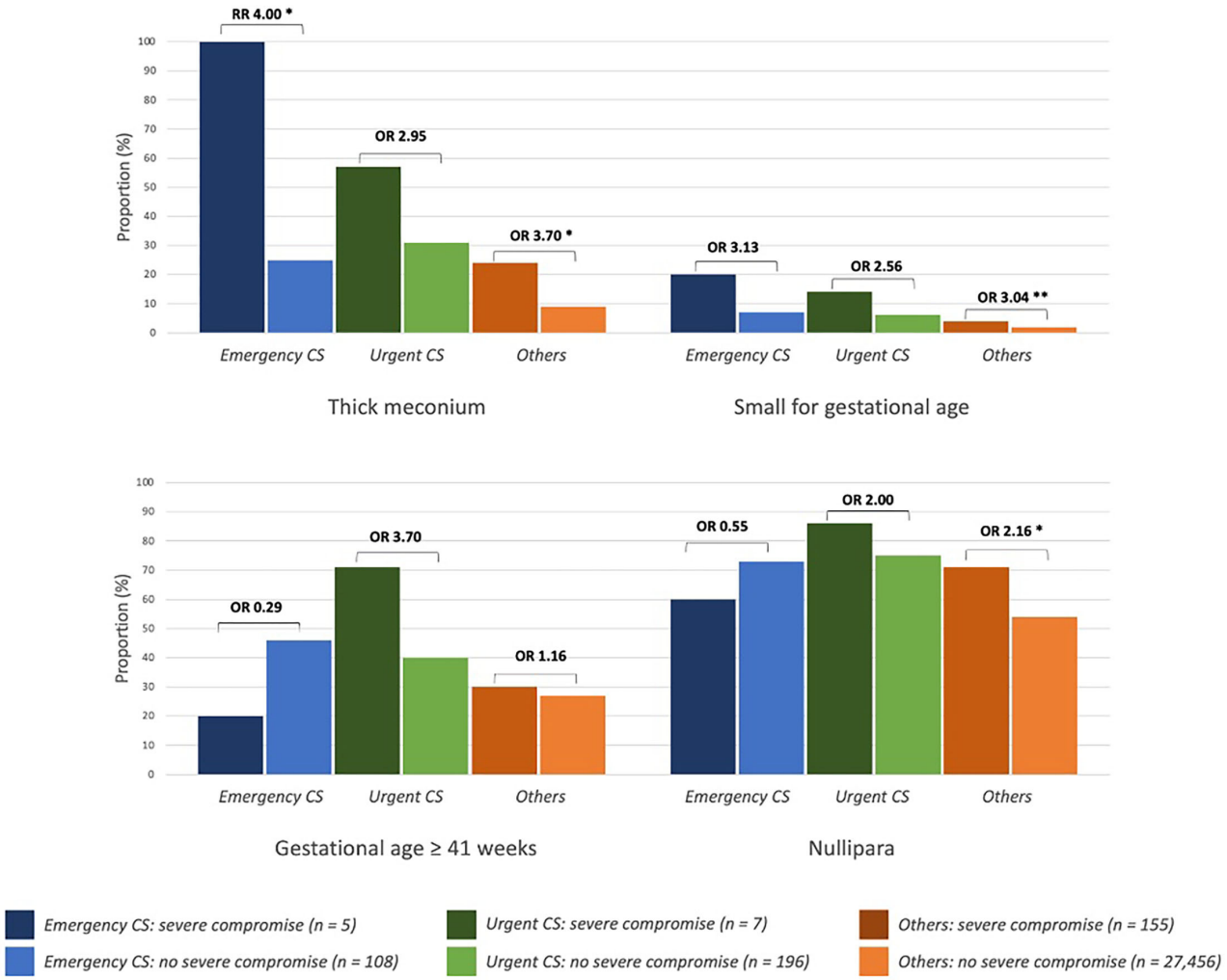

FIGURE 2 | Clinical characteristics of newborns with and without severe compromise. *Emergency CS*, emergency caesarean delivery (<2 h after start CTG monitoring); *Urgent CS*, urgent caesarean delivery (between 2 to 5 h after start CTG monitoring); RR, risk ratio; OR, odds ratio; $^{*}p \leq 0.01$, $^{**}p \leq 0.05$.

and hemorrhage due to birth injury; other birth injuries to central nervous system; birth injury to skeleton; and birth injury to peripheral nervous system.

Birth weight deviations were calculated with the adjusted Yudkin's chart percentiles (26) and reported in terms of small for gestational age (birth weights below the 3rd percentile) and large for gestational age (birth weights above the 97th percentile).

Severe compromise was defined as a composite outcome consisting of clinically relevant short-term perinatal outcomes collected at hospital level including stillbirth, neonatal death, seizures, neonatal encephalopathy, and need for intubation or intensive resuscitation followed by neonatal intensive care unit (NICU) admission for ≥48 h (15). Reported birth outcomes include *severe compromise*, Apgar score <4 at 1 min, Apgar score <7 at 5 min, and/or umbilical cord arterial pH <7.0 or <7.05 (27).

Onset of labor was defined by the attending clinician as per clinical practice (regular uterine contractions associated with cervical dilatation ≥3 cm) and documented in the routinely collected digital maternity notes.

Digital CTG data, sampled at 4 Hz as per the standard output of fetal monitors (typically Philips Avalon) were archived by a central monitoring system (Huntleigh Healthcare Ltd., Cardiff, UK). Computerized CTG features were calculated with the Oxford System (OxSys1.5) algorithms (15). OxSys analyses the CTG recordings in 15 min windows that move forward every 5 min; CTG characteristics are calculated for each window after heuristic noise removal algorithms. Pregnancy and labor data were derived from the Oxford Clinical Maternity Database, detailing maternal demographics, obstetric history, labor, delivery outcomes, and infant characteristics.

We considered only the first available 20–60 min of the CTG traces as part of the *initial assessment* (28), based on the recommended 20 min duration for CTG around labor onset (22) and additional time for the assessment of fetal behavioral states if they were absent in the first 20 min. Where available, the

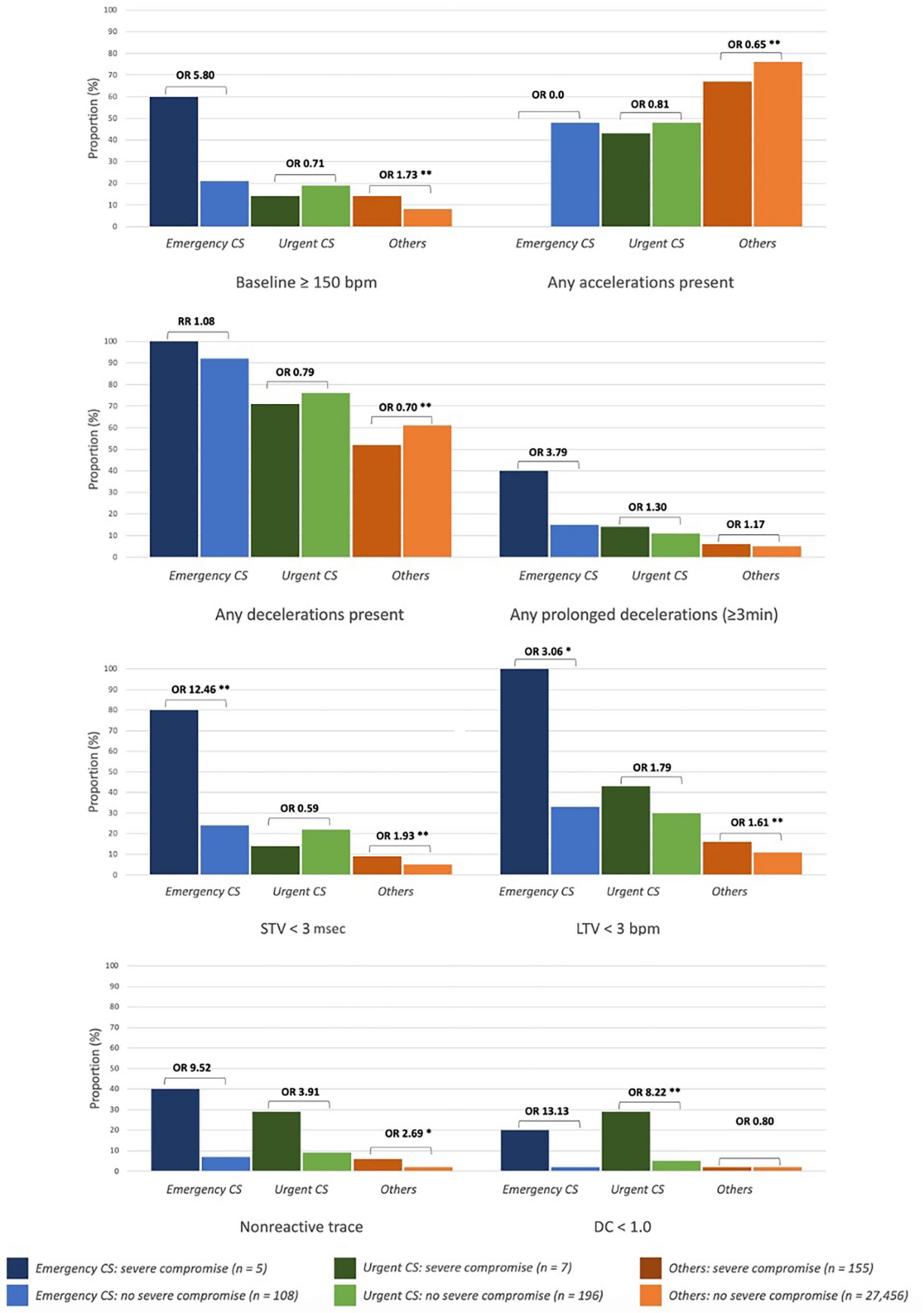

FIGURE 3 | First-hour CTG characteristics of newborns with and without severe compromise. *Emergency CS*, emergency caesarean delivery (< 2 h after start CTG monitoring); *Urgent CS*, urgent caesarean delivery (between 2 to 5 h after start CTG monitoring); DC, decelerative capacity; LTV, long-term variability; RR, risk ratio; OR, odds ratio. *$p \leq 0.01$, **$p \leq 0.05$.

entire first hour of CTG was analyzed and the results reported. Where not, the longest available CTG segment (≥20 min) was analyzed instead. CTG recordings made after the first hour were not analyzed in this study.

In addition to standard CTG characteristics such as baseline rate, variability, decelerations, and accelerations, OxSys1.5 computes the decelerative capacity (DC), analyzing the entire FHR signal in a moving 15 min window to provide an average measure of downward movements (29). Lower DC values are measured in a normal trace without decelerations, whereas increased DC values are measured in deep, steep-sloped, and/or repetitive decelerations (15). Non-reactive trace is defined as DC <1 bpm and no accelerations as previously reported (15).

Statistical Methods

Data were analyzed in MATLAB version R2019b (The MathWorks Inc., Massachusetts, USA). Medians with 25 and 75th percentiles were calculated for continuous and ordinal variables. Incidence with percentages were calculated for categorical variables. Univariate analysis of categorical variables was performed with the X^2 test and Fisher's exact test with relative risks, odds ratios, and 95% confidence intervals as appropriate. The Kruskal-Wallis one-way analysis of variance test was used to compare the medians and interquartile range (IQR) of continuous variables.

RESULTS

The CTG recording was commenced at different times in relation to the clinically determined time of the onset of active labor. As a result, 90% of all traces were within 10 h before and 6.5 h after labor onset (second stage recordings were excluded). The median difference between labor onset and CTG start was 7 min. The interquartile range is−135 and 230 min, negative values correspond to CTGs started after labor onset (46% of all), and positive values to those started before labor onset (56%).

Cesarean delivery for presumed fetal compromise was performed within 2 h of starting the CTG (*Emergency CS*) in 0.4% (113/27,927) and between 2 and 5 h (*Urgent CS*) in 0.7% (203/27,927) of all deliveries.

Majority of infants with severe compromise at birth were in *Others* ($n = 155$ at rate of 0.6%), with another five and seven in the *Emergency CS* and *Urgent CS* groups respectively (rates of 4.4 and 3.4%, respectively). **Figures 2**, **3** show the proportion of babies with selected characteristics of interest for the three groups (*Emergency CS*, *Urgent CS*, and *Others*), stratified for neonatal outcome. Dark colors correspond to the proportion of severely compromised babies and light colors to those without severe compromise.

Thick meconium and small for gestational age were more common in deliveries with severely compromised neonates across all groups; nulliparous women were more common in *Others* with severe compromise vs. *Others* without severe compromise (**Figure 2**). Post-date deliveries, induced labor, and maternal fever were more common clinical characteristics of severely compromised babies in the *Urgent CS* group and *Others*.

Figure 3 shows the first-hour CTG characteristics of severely compromised babies compared to those without severe compromise. Baseline ≥150 bpm, non-reactive trace, reduced long-term variability (LTV), and fewer accelerations were more prevalent in severely compromised neonates in all groups. Baseline FHR ≥150 bpm was more common in severely compromised babies in the *Emergency CS* group and *Others*. Reduced DC <1 bpm was significantly more common in severely compromised newborns with *Urgent CS*.

Selected details of clinical and CTG characteristics are shown in **Table 1** and the full details in the **Supplementary Table 1**.

Table 2 details clinical and CTG characteristics specifically for the neonates with severe compromise at birth ($n = 167$), and full details are in the **Supplementary Table 2**. These represent 4.4% (5/113) of babies with *Emergency CS*; 3.4% (7/203) of babies with *Urgent CS*; and 0.6% (155/27,611) of *Others*.

All severely compromised neonates in the *Emergency CS* group ($n = 5$) had thick meconium, meconium aspiration syndrome, and were admitted to the neonatal intensive care unit (NICU), and all of these were less common in the severely compromised infants in the other two groups. Their first-hour CTG traces were marked by baseline ≥150 and ≥160 bpm; reduced variability (both short-term variability (STV) and LTV); non-reactive trace; prolonged decelerations; and lack of accelerations. The first-hour CTGs are shown in **Figure 4**.

Severely compromised babies in the *Urgent CS* group ($n = 7$) were more often induced post-date labors in first-time mothers (**Table 2**) and had depressed Apgar scores and critically low umbilical cord arterial pH (group median pH was 6.83). Their first-hour CTG traces were marked by reduced variability (particularly LTV) and non-reactive trace, and the CTGs are shown in **Figure 5**.

Severely compromised neonates in *Others* ($n = 155$) were delivered by spontaneous vaginal delivery in 47.7% (74/155), instrumental vaginal delivery in 40% (62/155), and by cesarean section in 12.3% (19/155). Their first-hour CTG traces were marked by reduced variability (both LTV and STV) and non-reactive traces. A random selection of eight of these first-hour CTGs is shown in **Figure 6**.

DISCUSSION

Main Findings

Analysis of this routinely collected cohort of over 27,000 term births showed that an important proportion of fetuses with severe compromise or at risk of severe compromise had significantly different features of the first-hour CTG and presence of clinical risk factors such as thick meconium, small for gestational age, and maternal fever ≥38C. We considered three exclusive groups: infants delivered by CS for "presumed fetal compromise" within 2 h of starting the CTG (*Emergency CS*); between 2 and 5 h of starting the CTG (*Urgent CS*); and the rest of deliveries (*Others*). This data stratification allowed us to account for any effects the emergency or urgent cesarean sections had on labor outcomes, as well as to obtain estimates for the proportion of severely compromised infants for whom potential early signs were present. Severely compromised infants in these groups

TABLE 1 | Clinical and CTG characteristics of total cohort ($n = 29{,}927$).

	Emergency cesarean section, *Emergency CS* ($n = 113$)		**Urgent cesarean section, *Urgent CS* ($n = 203$)**		**All *Others* ($n = 27{,}611$)**		
	n* or *median	**% or *IQR***	***n* or *median***	**% or *IQR***	***n* or *median***	**% or *IQR***	***p*-value**
LABOR DETAILS							
Labor onset							
Established labor at start CTG	103	91.2	135	66.5	13,485	48.8	≤0.01
Antepartum risk factors							
Nulliparity	82	72.6	153	75.4	14,901	54.0	≤0.01
Post term (≥41 weeks)	51	45.1	84	41.4	7,522	27.2	≤0.01
Intrapartum risk factors							
Maternal fever (≥38.0°C)[†]	12[A]	12.4[A]	20[B]	11.2[B]	1,482[C]	5.6[C]	≤0.01
Thick meconium	32	28.3	65	32.0	2,182	7.9	≤0.01
Sentinel event	6	5.3	8	3.9	130	0.5	≤0.01
Delivery mode							
Cesarean section	113	100.0	203	100.0	798	2.9	–
Instrumental vaginal	–	–	–	–	8,044	29.1	–
Spontaneous vaginal	–	–	–	–	18,769	68.0	–
DELIVERY OUTCOME							
Objective fetal compromise							
Severe compromise[‡]	5	4.4	7	3.4	155	0.6	≤0.01
Resuscitation	4	3.5	9	4.4	199	0.7	≤0.01
Apgar score < 4 at 1 min	9	8.0	24	11.8	591[D]	2.1[D]	≤0.01
Apgar score < 7 at 5 min	3	2.7	8	3.9	229[E]	0.8[E]	≤0.01
Arterial umbilical cord pH	*7.23*[F]	*7.15–7.28*[F]	*7.22* [F]	*7.12–7.27*[F]	*7.22*[G]	*7.14–7.28*[G]	≤0.01
pH < 7.00	6[F]	2.8[F]	9[F]	4.6[F]	190[G]	1.0[G]	≤0.01
pH < 7.05	9[F]	8.4[F]	23[F]	11.7[F]	502[G]	2.5[G]	≤0.01
Mortality							
Stillbirth	0	0.0	0	0.0	0	0.0	–
Neonatal death	1	0.9	0	0.0	16	0.1	≤0.01
Morbidity							
Convulsions	3	2.7	1	0.5	43	0.2	≤0.01
Meconium aspiration syndrome	9	8.0	7	3.4	78	0.3	≤0.01
NICU admission	12	10.6	32	15.8	1,133	4.1[E]	≤0.01
Length of stay (days)	*6*	*2–10*	*4*	*2–7*	3	*1–5*	0.109
COMPUTERIZED CTG FEATURES (FIRST HOUR)							
Baseline (bpm)	*138*[H]	*129–148*[H]	*139*[I]	*130–148*[I]	*135*[J]	*127–142*[J]	≤0.01
≥150 bpm	25[H]	22.1[H]	38[I]	18.9[I]	2,299[J]	8.4[J]	≤0.01
≥160 bpm	14[H]	12.5[H]	11[I]	5.5[I]	585[J]	2.1[J]	≤0.01
Short-term variability	*5.4*[H]	*2.9–9.4*[H]	*4.6*[I]	*3.1–6.9*[I]	*6.1*[J]	*4.5–7.9*[J]	≤0.01
<3 msec	30[H]	26.8[H]	44[I]	21.9[I]	1,358[J]	5.0[J]	≤0.01
Long-term variability	*3.9*[H]	*2.5–6.3*[H]	*3.7*[K]	*2.8–4.9*[K]	*4.9*[L]	*3.8–6.2*[L]	≤0.01
<3 bpm	40[H]	35.7[H]	60[K]	30.0[K]	2,961[L]	10.9[L]	≤0.01
Non-reactive trace	9[H]	8.0[H]	20[I]	10.0[I]	624[M]	2.3[M]	≤0.01
Any accelerations present	51[H]	45.5[H]	96[I]	47.8[I]	20,572[M]	75.5[M]	≤0.01
Any decelerations present	103[I]	92.8[I]	150[N]	75.8[N]	16,2431[O]	61.29[O]	≤0.01
Any prolonged decelerations (≥3 min)	18[H]	16.0[H]	23[I]	11.4[I]	1,388[M]	5.1[M]	≤0.01
Decelerative capacity (bpm)	*3.2*[H]	*2.0–4.6*[H]	*2.3*[I]	*1.6–3.2*[I]	*2.5*[J]	*1.9–3.3*[J]	≤0.01
<1.0 bpm	3[H]	2.7[H]	11[I]	5.5[I]	664[J]	2.5[J]	≤0.01

n, number; IQR, inter-quartile range (25–75th percentiles); NICU, neonatal intensive care unit; bpm, beats per minute. Super-indices A–O indicate the number of missing data: A, 16; B, 24; C, 1074; D, 8; E, 11; F, 6; G, 7811; H, 1; I, 2; J, 362; K, 3; L, 388; M, 359; N, 3; O, 752. [†]Maternal fever defined as one-time maximum maternal temperature measurement of ≥38.0°C, based on UK maternity guidelines (29). [‡]Composite outcome: stillbirth, neonatal death, seizures, neonatal encephalopathy, intubation or resuscitation followed by NICU admission for ≥ 48 h.

TABLE 2 | Clinical and CTG characteristics of severely compromised newborns ($n = 167$).

	Emergency cesarean section, *Emergency CS* ($n = 5$)		Urgent cesarean section, *Urgent CS* ($n = 7$)		All *Others* ($n = 155$)		
	n* or *median	**% or *IQR***	***n* or *median***	**% or *IQR***	***n* or *median***	**% or *IQR***	***p*-value**
LABOR DETAILS							
Labor onset							
Established labor at start CTG	5	100.0	5	71.4	54	34.8	≤0.01
Antepartum risk factors							
Nulliparity	3	60.0	6	85.7	111	71.6	0.601
Post term (≥41 weeks)	1	20.0	5	71.4	47	30.2	0.062
Intrapartum risk factors							
Maternal fever (≥38.0°C)[†]	0	0.0	1[A]	16.7[A]	13[B]	8.6[B]	0.618
Thick meconium	5	100.0	4	57.1	37	23.9	≤0.01
Sentinel event	0	0.0	0	0.0	1	0.6	0.962
Delivery mode							
Cesarean section	5	100.0	7	100.0	19	12.3	–
Instrumental vaginal	–	–	–	–	62	40.0	–
Spontaneous vaginal	–	–	–	–	74	47.7	–
DELIVERY OUTCOME							
Objective fetal compromise							
Severe compromise[‡]	5	100.0	7	100.0	155	100.0	–
Resuscitation	3	60.0	6	85.7	98	63.2	0.470
Apgar score <4 at 1 min	1	20.0	6	85.7	85[A]	55.2[A]	0.076
Apgar score <7 at 5 min	1	20.0	4	57.1	75[C]	49.0[C]	0.396
Arterial umbilical cord pH	*7.14*	*7.09–7.28*	*6.83*[A]	*6.78–7.01*[A]	*7.14*[D]	*7.05–7.23*[D]	≤0.01
pH < 7.00	0	0.0	4[A]	66.7[A]	26[D]	19.3[D]	≤0.01
pH < 7.05	1	20.0	5[A]	88.3[A]	33[D]	24.4[D]	≤0.01
Mortality							
Stillbirth	0	0.0	0	0.0	0	0.0	–
Neonatal death	1	20.0	0	0.0	14	9.0	0.488
Morbidity							
Convulsions	3	60.0	1	14.3	43	27.7	0.203
Meconium aspiration syndrome	5	100.0	1	14.3	19	12.3	≤0.01
NICU admission	5	100.0	6	85.7	131	84.5	0.633
Length of stay (days)	*9*	*8–18*	*5*	*4–6*	***3***	*3–8*	0.168
COMPUTERIZED CTG FEATURES (FIRST HOUR)							
Baseline (bpm)	*158*	*145–164*	*135*	*127–145*	*136*[C]	*129–144*[C]	0.014
≥150 bpm	3	60.0	1	14.3	21[C]	13.7[C]	0.027
≥160 bpm	2	40.0	0	0.0	8[C]	5.2[C]	≤0.01
Short-term variability	*1.8*	*1.6–2.7*	5.6	*3.2–7.0*	*5.3*[C]	*4.1–7.8*[C]	≤0.01
<3 msec	4	80.0	1	14.3	14[C]	9.2[C]	≤0.01
Long-term variability	*2.1*	*1.6–2.6*	*4.4*	*2.1–5.2*	*4.6*[C]	*3.4–6.1*[C]	≤0.01
<3 bpm	5	100.0	3	42.9	25[C]	16.3[C]	≤0.01
Non-reactive trace	2	40.0	2	28.6	9[C]	5.9[C]	≤0.01
Any accelerations present	0	0.0	3	42.9	102[C]	66.7[C]	0.013
Any decelerations present	5	100.0	5	71.4	80[A]	52.6[A]	0.125
Any prolonged decelerations (≥3 min)	2	40.0	1	14.3	9[C]	5.8[C]	0.012
Decelerative capacity (bpm)	*1.8*	*1.0–2.4*	*2.3*	*1.2–3.8*	*2.4*[C]	*1.7–3.1*[C]	0.301
<1.0 bpm	1	20.0	2	28.6	3[C]	2.0[C]	≤0.01

n, number; IQR, inter-quartile range (25–75th percentiles); NICU, neonatal intensive care unit; bpm, beats per minute. Super-indices A–D indicate the number of missing data: A, 1; B, 4; C, 2; D, 20. [†]Maternal fever defined as one-time maximum maternal temperature measurement of ≥38.0°C, based on UK maternity guidelines (29). [‡]Composite outcome: stillbirth, neonatal death, seizures, neonatal encephalopathy, intubation or resuscitation followed by NICU admission for ≥ 48 h.

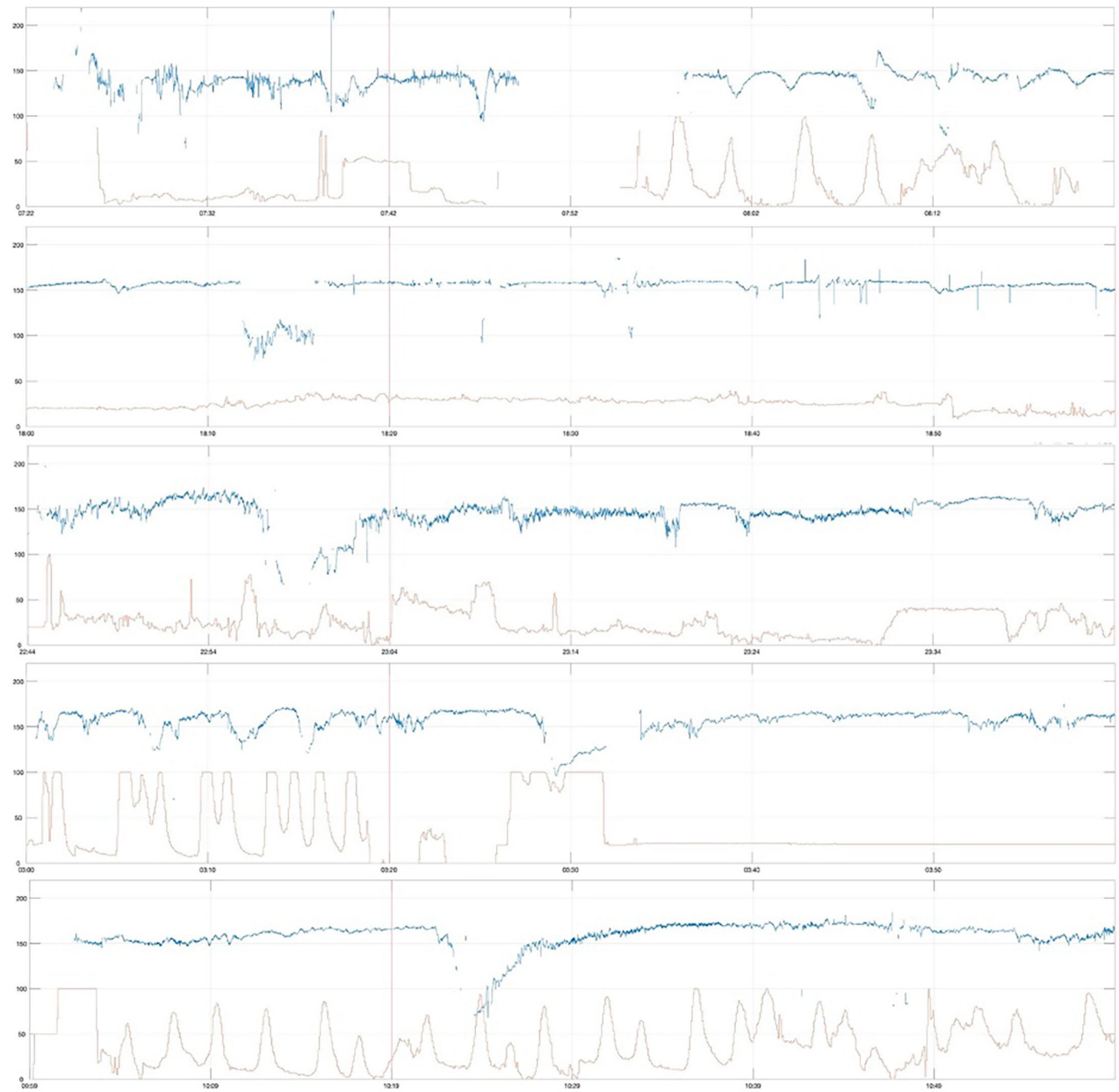

FIGURE 4 | First-hour CTG of severely compromised newborns delivered by cesarean section within 2 h after CTG start (emergency cesarean delivery, *Emergency CS*), vertical red line at 20 min, paper speed 1 cm/min, vertical scale 20 beats per min/cm.

had higher rates of baseline FHR ≥150 bpm, non-reactive initial CTG, reduced short-term and long-term variability, less accelerations, and decelerative capacity amongst fetuses with severe compromise. Prolonged decelerations (≥3 min) were also more common in the first hour CTG of infants born with severe compromise. None of the severely compromised infants in the *Urgent CS* had a first hour baseline ≥160 bpm and only a small proportion had STV < 3 msec. In this group, non-reactive initial trace and DC < 1 bpm were stronger risk factors in about a third of severely compromised infants. There was a significantly higher proportion of babies with arterial cord pH < 7.00 in the severely compromised babies with *Urgent CS* (**Table 2**), with the length of stay in NICU longest for the severe compromised babies in the *Emergency CS* group, suggesting that the components of the "severe compromise" varied between the groups. Umbilical artery acidemia was absent in the majority of severely compromised infants in *Emegency CS* and *Others*. The rate of decelerations in the initial CTG was not

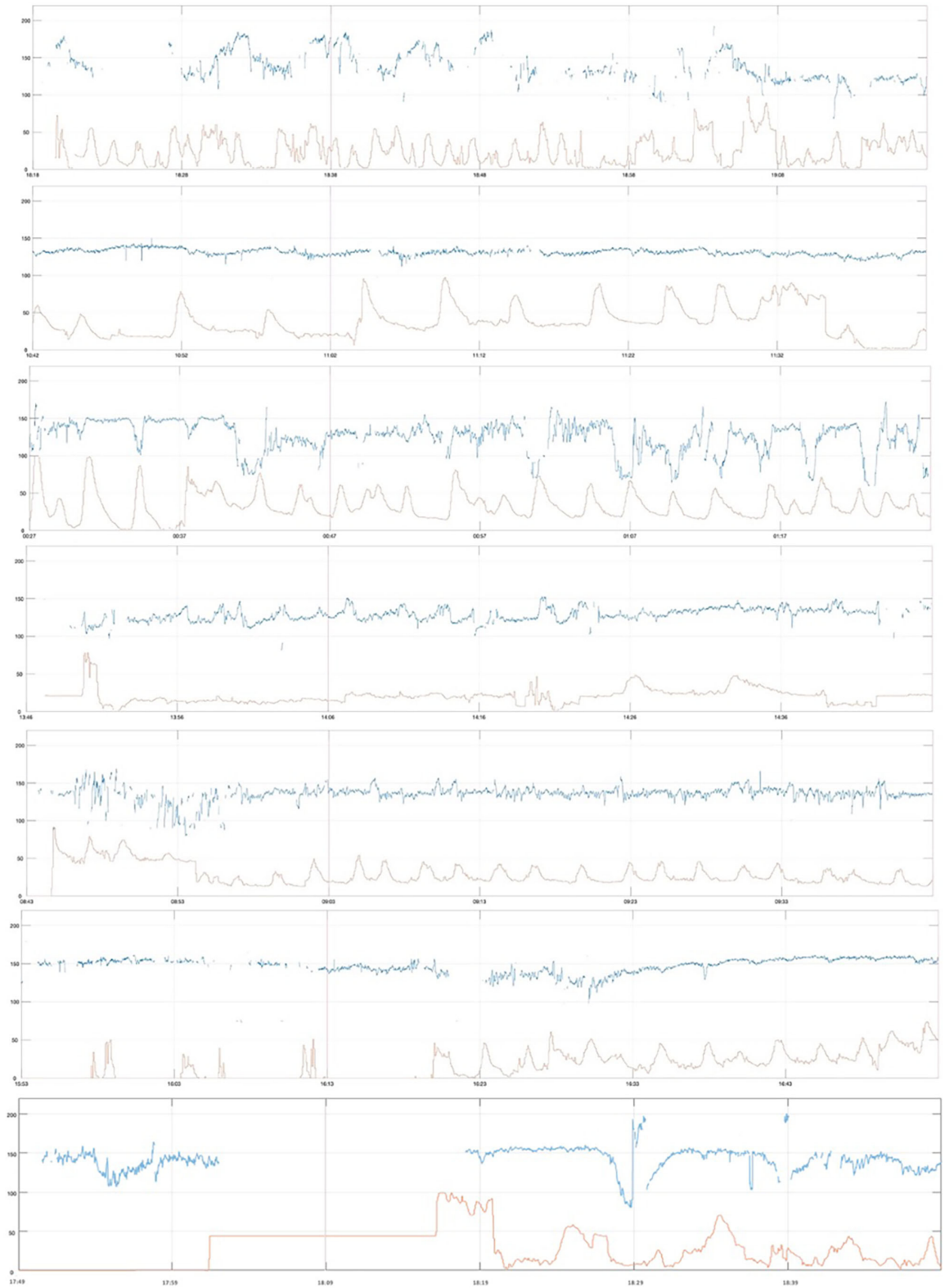

FIGURE 5 | First-hour CTG of severely compromised newborns delivered by cesarean section between 2 and 5 h after start of CTG (urgent cesarean delivery, *Urgent CS*), vertical red line at 20 min, paper speed 1 cm/min, vertical scale 20 beats per min/cm.

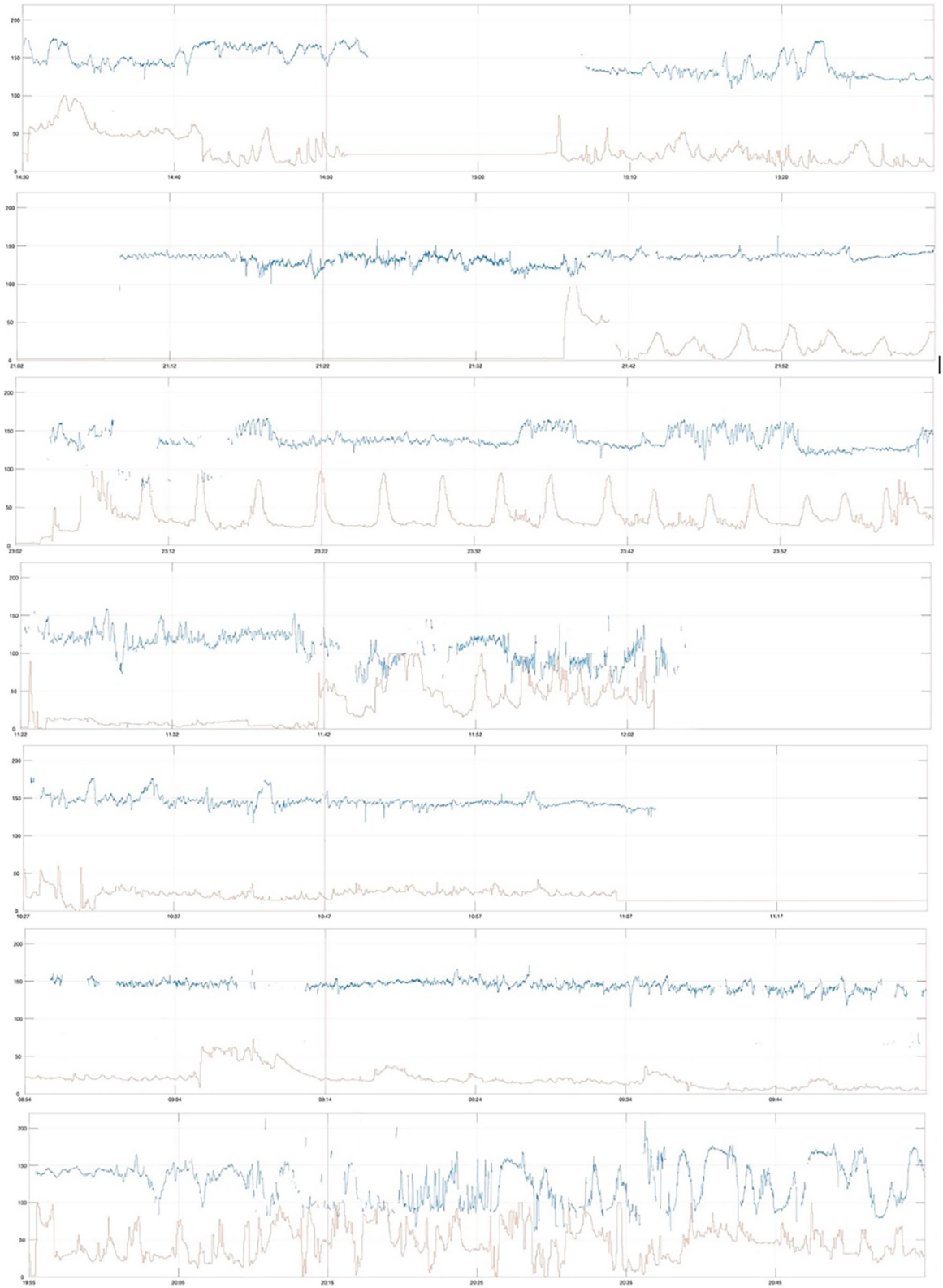

FIGURE 6 | First-hour CTG traces of severely compromised newborns of the remaining cohort (*Others*), random selection of 8 traces, vertical red line at 20 min, paper speed 1 cm/min, vertical scale 20 beats per min/cm.

associated with severe compromise, but our data was "censored" by clinical practice.

In addition, we demonstrated that computerized analysis of the CTG adds value by permitting the analysis of large cohorts to support clinical guidelines.

Clinical Implications and Interpretation

Our data suggest that clinical guidelines for CTG interpretation for the *initial* trace may need to differ from those applied to the rest of labor. Perhaps there should be less emphasis on the presence of decelerations *per se*, unless they are prolonged (≥3 min). Instead, there should be a lower threshold for the definition of FHR "tachycardia" at this time (i.e., 150 bpm), an assessment for a non-reactive trace (i.e., lack of cyclicity or reactivity). Our data also confirm the relevance of and provide evidence for STV and LTV, as well as clinical risk factors such as thick meconium, small for gestational age, and others.

These findings suggest that a consistent assessment of the CTG together with known clinical risk factors around the onset of labor has the potential to identify early a small but important proportion of babies at risk of severe compromise. Our data shows that it is likely that some of the cases were "missed" in the group of *Others* and recognized late in the *Urgent CS* group. Some of the CTG characteristics important for the first-hour CTG assessment, such as reduced variability, decelerative capacity, and the non-reactive trace, are not typically detectable with intermittent auscultation (20).

The primary outcome of "severe compromise" was broadly defined to include cases which may not have their origins in hypoxic-ischaemic insults. Approximately half of small for gestational age fetuses attributable to placental insufficiency are missed in the antenatal period, and about 10% of them sustain severe compromise (30). Other vulnerabilities, including intrauterine infection, cerebral hemorrhage, and maternal hyperglycaemia (13, 18, 19, 31), may also alter fetal physiological adaptations to reduced oxygen during labor (18, 19) in a manner that is difficult to detect with conventional CTG (13, 18, 31). Therefore, in principle, clinical factors and fetal physiology should be considered for accurate interpretation of CTG in the initial assessment of women in labor. Our findings confirm the importance of both first-hour CTG and clinical characteristics.

Our findings are relevant for fetuses for whom cesarean delivery before or in early labor may avoid the more dramatic emergency deliveries and the attendant maternal and fetal risks (30). Furthermore, better risk assessments with CTG in the initial assessment could enhance individualized care, informed decision-making, and improve maternal engagement during labor (17, 32).

Strengths and Limitations

Using computerized methods, we studied a large, detailed, and complete, high-quality maternity database. Specific markers of brain development were not collected at the hospital level and therefore unavailable for our data analysis of routinely collected cohort data. We included all eligible infants resulting in nearly equal proportion of babies with CTG starting before and after the onset of labor, 46 vs. 54% respectively. Therefore, our findings are relevant to both periods, pre and early labor, largely around onset of labor where women are assessed in a maternity unit triage room or bays. This is particularly important because the definition of onset of labor and CTG initiation varies and depends, amongst other things, on the time the woman presents to the hospital.

Clinical practice has inevitably changed since the timespan of this dataset and the data for the subsequent 10-year period (2011–2020) were not fully available yet at the time of analysis, but their analysis will follow the presented format in this study to allow comparison.

A limitation of this study was the lack of information for the cervical dilatation at the time of initial CTG recording as well as the fact that our routinely collected data relate to a high-risk obstetric population, as per clinical CTG use in the UK.

We propose to build on the various univariate analyses presented in the current study to undertake multivariate analyses in future studies using larger datasets.

CONCLUSION

This study reports that a small but important proportion of infants born with severe compromise had significantly different computerized CTG characteristics around the onset of labor (detected with OxSys algorithms), typically in the context of other clinical risk factors. Clinical guidelines for CTG interpretation, for the *initial* trace specifically, may need to be different from those for monitoring throughout labor, with less focus on the presence of decelerations *per se*, unless they are prolonged (≥3 min). Instead, there should be a lower threshold for the definition of FHR "tachycardia" and assessment for non-reactive trace (lack of cyclicity or reactivity). We further confirm the relevance of and provide evidence for STV and LTV, as well as clinical risk factors including thick meconium and small for gestational age.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The Ethics Approval does not allow for sharing the data publicly. Requests to access these datasets should be directed to antoniya.georgieva@wrh.ox.ac.uk.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Newcastle & North Tyneside 1 Research Ethics Committee, Reference 11/NE0044 (data before 2008), and from the South Central Ethics Committee, Reference 13/SC/0153 (for data beyond 2008). Written informed consent for participation was not provided by the participants' legal guardians/next of kin because: This was routinely collected and

fully anonymised clinical data, and no personal information was processed.

AUTHOR CONTRIBUTIONS

AG conceived the study, pre-processed the data, and analyzed the CTGs. AL analyzed the full data, conducted the literature review, and wrote the first manuscript draft. AU provided clinical oversight and input to the research question, data analysis, and manuscript. All authors reviewed and contributed to writing the manuscript.

FUNDING

AG was funded by the UK National Institute of Health Research (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.

ACKNOWLEDGMENTS

We are indebted to Pro. Christopher WG Redman, Mr. Pawel Szafranski, and Ms. Mary Moulden for their substantial and invaluable work to collect and curate the database, analyzed in the present study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2022.784439/full#supplementary-material

REFERENCES

1. Lawn JE, Blencowe H, Oza S, You D, Lee ACC, Waiswa P, et al. Every newborn: progress, priorities, and potential beyond survival. *Lancet.* (2014) 384:189–205. doi: 10.1016/S0140-6736(14)60496-7
2. Lee ACC, Kozuki N, Blencowe H, Vos T, Bahalim A, Darmstadt GL, et al. Intrapartum-related neonatal encephalopathy incidence and impairment at regional and global levels for 2010 with trends from 1990. *Pediatr Res.* (2013) 74:50–72. doi: 10.1038/pr.2013.206
3. Akseer N, Lawn JE, Keenan W, Konstantopoulos A, Cooper P, Ismail Z, et al. Ending preventable newborn deaths in a generation. *Int J Gynecol Obstet.* (2015) 131:S43–8. doi: 10.1016/j.ijgo.2015.03.017
4. Zeitlin J, Mortensen L, Cuttini M, Lack N, Nijhuis J, Haidinger G, et al. Declines in stillbirth and neonatal mortality rates in Europe between 2004 and 2010: Results from the Euro-Peristat project. *J Epidemiol Community Health.* (2016) 70:609–15. doi: 10.1136/jech-2015-207013
5. De Bernis L, Kinney MV, Stones W, Ten Hoope-Bender P, Vivio D, Leisher SH, et al. Stillbirths: ending preventable deaths by 2030. *Lancet*. (2016) 387:703–16. doi: 10.1016/S0140-6736(15)00954-X
6. Rowe R, Draper ES, Kenyon S, Bevan C, Dickens J, Forrester M, et al. Intrapartum-related perinatal deaths in births planned in midwifery-led settings in Great Britain: findings and recommendations from the ESMiE confidential enquiry. *BJOG.* (2020) 127:1665–75. doi: 10.1111/1471-0528.16327
7. Joyce NM, Tully E, Kirkham C, Dicker P, Breathnach FM. Perinatal mortality or severe neonatal encephalopathy among normally formed singleton pregnancies according to obstetric risk status:" is low risk the new high risk?" a population-based cohort study. *Eur J Obstet Gynecol Reprod Biol.* (2018) 228:71–5. doi: 10.1016/j.ejogrb.2018.06.010
8. Prior T, Mullins E, Bennett P, Kumar S. Prediction of fetal compromise in labor. *Obstet Gynecol.* (2014) 123:1263–71. doi: 10.1097/AOG.0000000000000292
9. Ayres-de-Campos D. Electronic fetal monitoring or cardiotocography, 50 years later: what's in a name? *Am J Obstet Gynecol.* (2018) 218:545–6. doi: 10.1016/j.ajog.2018.03.011
10. Alfirevic Z, Devane D, Gyte GM, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane database Syst Rev.* (2017) 2:CD006066. doi: 10.1002/14651858.CD006066.pub3
11. Georgieva A, Abry P, Chudacek V, Djuric PM, Frasch MG, Kok R, et al. Computer-based intrapartum fetal monitoring and beyond: a review of the 2nd workshop on signal processing and monitoring in labor (October 2017, Oxford, UK). *Acta Obstet Gynecol Scand.* (2019) 98:1207–17. doi: 10.1111/aogs.13639
12. Schifrin BS, Soliman M, Koos B. Litigation related to intrapartum fetal surveillance. *Best Pract Res Clin Obstet Gynaecol.* (2016) 30:87–97. doi: 10.1016/j.bpobgyn.2015.06.007
13. 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) 121:1063–70. doi: 10.1111/1471-0528.12900
14. NHS England. *Better Births. Improving Outcomes Of Maternity Services In England.* National Maternity Review (2016). p. 1–125.
15. Georgieva A, Redman CWG, Papageorghiou AT. Computerized data-driven interpretation of the intrapartum cardiotocogram: a cohort study. *Acta Obstet Gynecol Scand.* (2017) 96:883–91. doi: 10.1111/aogs.13136
16. Institute of Medicine and National Research Council. *An Update on Research Issues in the Assessment of Birth Settings. Workshop Summary*. Washington, DC: The National Academies Press (US) (2013). p. 3.
17. Devane D, Lalor JG, Daly S, McGuire W, Cuthbert A, Smith V. Cardiotocography versus intermittent auscultation of fetal heart on admission to labour ward for assessment of fetal wellbeing. *Cochrane Database Syst Rev*. (2017) 1:CD005122. doi: 10.1002/14651858.CD005122.pub5
18. Lear CA, Westgate JA, Ugwumadu A, Nijhuis JG, Stone PR, Georgieva A, et al. Understanding fetal heart rate patterns that may predict antenatal and intrapartum neural injury. *Semin Pediatr Neurol.* (2018) 28:3–16. doi: 10.1016/j.spen.2018.05.002
19. Turner JM, Mitchell MD, Kumar SS. The physiology of intrapartum fetal compromise at term. *Am J Obstet Gynecol.* (2020) 222:17–26. doi: 10.1016/j.ajog.2019.07.032
20. Parts L, Holzmann M, Norman M, Lindqvist PG. Admission cardiotocography: a hospital based validation study. *Eur J Obstet Gynecol Reprod Biol.* (2018) 229:26–31. doi: 10.1016/j.ejogrb.2018.07.016
21. Blix E. The admission CTG: is there any evidence for still using the test? *Acta Obstet Gynecol Scand.* (2013) 92:613–9. doi: 10.1111/aogs.12091
22. Ingemarsson I, Arulkumaran S, Ingemarsson E, Tambyraja RL, Ratnam SS. Admission test: a screening test for fetal distress in labor. *Obstet Gynecol.* (1986) 68:800–6.
23. Turnbull D, Salter A, Simpson B, Mol BW, Chandraharan E, McPhee A, et al. Comparing the effect of STan (cardiotocographic electronic fetal monitoring (CTG) plus analysis of the ST segment of the fetal electrocardiogram) with CTG alone on emergency caesarean section rates: study protocol for the STan Australian randomised controlled trial (START). *Trials.* (2019) 20:539. doi: 10.1186/s13063-019-3640-9
24. Blix E, Brurberg KG, Reierth E, Reinar LM, Øian P. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: a systematic review and meta-analysis of randomized trials. *Acta Obstet Gynecl Scand.* (2016) 95:16–27. doi: 10.1111/aogs.12828

25. INFANT Collaborative Group. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet.* (2017) 389:1719–29. doi: 10.1016/S0140-6736(17)30568-8
26. Yudkin PL, Aboualfa M, Eyre JA, Redman CWG, Wilkinson AR. New birthweight and head circumference centiles for gestational ages 24 to 42 weeks. *Early Hum Dev.* (1987) 15:45–52. doi: 10.1016/0378-3782(87)90099-5
27. Ayres-de-Campos D, Arulkumaran S. FIGO consensus guidelines on intrapartum fetal monitoring: physiology of fetal oxygenation and the main goals of intrapartum fetal monitoring. *Int J Gynaecol Obstet.* (2015) 131:5–8. doi: 10.1016/j.ijgo.2015.06.018
28. National Institute for Health and Care Excellence (NICE). *Intrapartum Care For Healthy Women And Babies. Clinical guideline [CG190].* (2014). Available online at: https://www.nice.org.uk/guidance/cg190 (accessed August 14, 2020).
29. 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) 121:889–94. doi: 10.1111/1471-0528.12568
30. Lindqvist PG, Molin J. Does antenatal identification of small-for-gestational age fetuses significantly improve their outcome? *Ultrasound Obstet Gynecol.* (2005) 25:258–64. doi: 10.1002/uog.1806
31. Dhillon SK, Lear CA, Galinsky R, Wassink G, Davidson JO, Juul S, et al. The fetus at the tipping point: modifying the outcome of fetal asphyxia. *J Physiol.* (2018) 596:5571–92. doi: 10.1113/JP274949
32. Mendis R, Flatley C, Kumar S. Maternal demographic factors associated with emergency caesarean section for non-reassuring foetal status. *J Perinat Med.* (2018) 46:641–7. doi: 10.1515/jpm-2017-0142

Author Disclaimer: The views expressed here are those of the authors and not necessarily those of the NIHR or the Department of Health (UK).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership