U.PORTO

FMUP FACULDADE DE MEDICINA
UNIVERSIDADE DO PORTO

COMPLEXITY SCIENCES APPLIED TO CARDIOTOCOGRAPHY

João Manuel Monteiro dos Santos

Tese de Doutoramento apresentada à Faculdade de Medicina da Universidade do Porto em Investigação Clínica e Serviços de Saúde

Porto, 2020

COMPLEXITY SCIENCES APPLIED TO CARDIOTOCOGRAPHY

João Manuel Monteiro dos Santos

Orientadora:
Prof. Cristina Santos
MEDCIDS – Departamento de Medicina da Comunidade, Informação e Decisão em Saúde,
Faculdade de Medicina da Universidade do Porto
CINTESIS – Centro de Investigação em Tecnologias e Serviços de Saúde,
Faculdade de Medicina da Universidade do Porto

Co-Orientadores:
Prof. Luís Antunes
CRACS – Center for Research in Advanced Computing Systems,
Faculdade de Ciências da Universidade do Porto
Departamento de Ciência de Computadores da Faculdade de Ciências da Universidade do Porto

Prof. Teresa Henriques
MEDCIDS – Departamento de Medicina da Comunidade, Informação e Decisão em Saúde,
Faculdade de Medicina da Universidade do Porto
CINTESIS – Centro de Investigação em Tecnologias e Serviços de Saúde,
Faculdade de Medicina da Universidade do Porto

Porto, 2020

Acknowledgements

Greatest considerations go to my supervisors Cristina Santos and Luís Antunes provided me all the tools and opportunities I could have ever wished throughout my academic paths, both in Masters and PhD degrees. They provided me a great atmosphere for the development of my thesis. Teresa Henriques was also essential for the success of this dissertation. Her feedback and guidance had a great positive impact in this work, whose support and perseverance were key factors.

I would like to leave a note of appreciation to Andreia Teixeira, Luísa Castro, Rosário Ribeiro, Mohammad Nozari and Marcelo Santos from Department of Community Medicine, Information and Health Decision Sciences – MEDCIDS, with whom I met frequently to discuss research topics. I also acknowledge members of SisPorto, namely João Bernardes, Hernani Gonçalves and Inês Nunes. Their direct contribution to the research papers and great atmosphere during all these years will for sure have a special spot in my heart.

To my colleagues at the Center for Research in Technology and Health Services - CINTESIS, professors at MEDCIDS and doctoral colleagues, a big thanks for all the knowledge given and for maintaining a smile and a positive vibe in the workplace. Going to work knowing that I will find such a relaxed atmosphere, provided me moments of gratitude and pride just for being part of it.

I want to thank research fellows from the Digi-Newb project. It was such a pleasure to be involved in this project. I also want to leave a note to Joana Dumas from CRACS for assisting me in everything I needed regarding this project.

In a personal manner, I must give all credit to my parents. They always made sure I had everything they could offer me to reach success. Their great sacrifice towards me and my brother are the example I want to follow.

And for supporting me, a big thank you to my girlfriend, brother and grandparents.

To all, thank you very much.

Table of Contents

Abstract

Regulation of physiological systems governing the functioning of the human body is based on a complex adjustment of various variables according to internal requirements and external influences. This regulation occurs across different temporal scales. Although any physiological control mechanism operates along well assigned temporal scales, different regulatory mechanisms tend to partially share time scales. In addition, control mechanisms might interact with each other according to nonlinear relationships and to causality schemes, thus producing information transfer from a signal to another one and cross-influences characterized by time scales different from those of the original interacting mechanisms. The study of this "complex physiology" needs mathematical approaches capable of dealing with concepts of complexity, scaling behavior, information transfer, and causality.

A complex system behaves in a nonlinear fashion, i.e., small changes in the input may lead to great changes in the outcome. High levels of complexity have been associated with a low agreement and high uncertainty in the clinical practice and, in complex scenarios, uncritical adherence to clinical guidelines may do more harm than good. Other approaches to deal with uncertainty should be used, for instance, nonlinear models and mining of patterns. The agreement among clinicians in the diagnosis and decision is frequent and surprisingly poor.

Cardiotocography used to assess fetal well-being, by the record of the fetal heart rate and the uterine contractions, is a good example of a procedure used in clinical practice but with a low agreement among clinicians and a poor prediction of neonatal state.

This thesis aims to explore how complexity analysis relates to fetal heart rate variability and how it can be used to prevent bad neonatal outcomes, as pre-term birth and fetal asphyxia. It is an objective of this thesis to find ways to better predict bad neonatal outcome.

In **Chapter 1** the concepts and rationale inherent to the topic are introduced and the objectives of the dissertation are defined. Also, the theoretical background for linear and nonlinear assessment of fetal heart rate (FHR) is reviewed and a brief introduction to heart rate analysis is presented, in particular for human fetuses. This chapter ends with the definitions of spectral analysis and nonlinear measures utilized in this dissertation.

The **Chapter 2** comprises the outline of this thesis.

The association between nonlinear indexes (entropy and compression) and physiological characteristics of FHR tracings using the FIGO international guidelines for fetal monitoring are presented in **Chapter 3**, as well as the evidence of complementary information captured by both entropy and compression. Here, multiscale analysis is as well used, in particular for fetal acidemia prediction during intrapartum. The results suggest complementarity of the nonlinear indexes studied, consistency of measures to changes in parameters, and changes in behavior at different scales, suggesting different dynamics in the interaction between the central and autonomic nervous systems with a FHR at different scales.

Chapter 4 tackles a different issue in fetal assessment: labor prediction, and, therefore, prematurity. Labor prediction is usually based on physiological characteristics. This chapter explores the possible contribution of entropy and complexity in this field. The results suggest an improvement in the ability to predict labour at both one and two weeks when the model included compression.

Chapter 5 addresses a different perspective for the detection of fetal acidemia during labor, as it presents an algorithm based on spectral analysis.

The dissertation ends in **Chapter 6** with an overall discussion and main conclusions.

In the signal analysis, and particularly in heart rate variability analysis, complexity measures are in focus by the scientific community, not only due to the existing requirements by the signal characteristics but mainly due to the data processing and analysis capabilities that current technologies provide and by the amount of existing data.

The challenge presented is related to the difficult interpretation of nonlinear measures in a physiological context, making it challenging to integrate decision support systems in clinical practice. In an era where wearables are emerging, continuous monitoring will provide researchers with an enormous amount of information, enabling the development of multivariate models, such as "deep learning”.

Complexity Science is therefore of great importance in the prevention and prediction of abnormal labour outcome, aiding healthcare professionals in clinical decisions. This dissertation took a step to the current knowledge of how the science of complexity can be better understood in the clinical context used in the prevention of bad neonatal outcomes.

Resumo

A regulação de sistemas fisiológicos encarregues do funcionamento do corpo humano é baseado num ajustamento complexo de diversas variáveis de acordo com necessidades internas e estímulos externos. Esta regulação ocorre em diversas escalas temporais. Embora qualquer mecanismo de controlo fisiológico opere em escalas temporais bem definidas, diferentes mecanismos de regulação tendem a partilhar, parcialmente, escalas de tempo. Mais ainda, mecanismos de controlo podem interagir de acordo com modelos não-lineares e esquemas de causalidade, produzindo assim, transferências de informação de um sinal para o outro e influências cruzadas caracterizadas por diferentes escalas de tempo dos mecanismos de interação originais. O estudo da "complexidade fisiológica" necessita de abordagens matemáticas capazes de lidar com conceitos de complexidade, comportamento de escala, transferência de informação e causalidade.

Um sistema complexo comporta-se de uma forma não-linear, isto é, pequenas alterações podem levar a resultados completamente díspares. Níveis elevados de complexidade foram associados a baixos níveis de concordância e alta incerteza na prática clínica e, em cenários complexos, adesão não crítica a diretrizes poderá causar danos maiores. Outras abordagens deveriam ser usadas para lidar com a incerteza, tais como os modelos não-lineares e a descoberta de padrões. De facto, a concordância entre especialistas no diagnóstico e tomadas de decisão é surpreendentemente baixa.

Cardiotocografia, usada para avaliar o bem-estar fetal, através do batimento cardíaco fetal e contrações uterinas, é um bom exemplo de um procedimento na prática clínica com baixa concordância na análise entre os especialistas e uma pobre predição do bem-estar neonatal.

Esta dissertação tem como objetivos explorar em como a análise de complexidade se relaciona com a variabilidade de batimento cardíaco fetal e na forma como pode ser usada na prevenção de partos com desfechos não desejados, como partos prematuros ou asfixia fetal. Também é um objetivo desta tese encontrar formas de prever melhor maus desfechos neonatais.

No **Capítulo 1** são introduzidos os conceitos inerentes ao tema e definidos os objetivos da dissertação. Além disso, a base teórica para a avaliação linear e não linear da Frequência Cardíaca Fetal (FCF) é revista e é apresentada uma breve introdução à análise da frequência cardíaca, em particular para fetos humanos. Este capítulo termina com as definições de análise espectral e das medidas não lineares utilizadas nesta dissertação.

O **Capítulo 2** apresenta a estrutura da dissertação

A associação entre os índices não lineares entropia e compressão e as características fisiológicas dos traçados de FCF usando as recomendações internacionais da FIGO para monitorização fetal são apresentadas no **Capítulo 3**, bem como a evidência sobre a informação complementar capturada tanto pela entropia quanto pela compressão. Aqui, a análise multi-escala também é usada, em particular para a previsão da acidemia fetal durante o parto. Os resultados sugerem complementaridade dos índices não lineares estudados, consistência das medidas para mudanças nos parâmetros e mudanças no comportamento em diferentes escalas, sugerindo diferentes dinâmicas na interação entre os sistemas nervoso central e autónomo com uma FCF em diferentes escalas.

O **Capítulo 4** aborda uma questão diferente na avaliação fetal: a previsão do parto e, portanto, a prematuridade. Usualmente a previsão do parto é feita com base em características fisiológicas. Neste capítulo explora-se o contributo que a entropia e a complexidade podem dar para esta previsão do parto. Os resultados sugerem uma melhoria na capacidade de previsão do parto em uma e duas semanas quando o modelo incluiu compressão.

O **Capítulo 5** aborda uma perspetiva diferente para deteção de acidemia fetal durante o parto, pois apresenta um algoritmo baseado em análise espectral.

A dissertação termina no **Capítulo 6** com uma discussão geral e as principais conclusões.

Na análise de sinal, e em particular em batimento cardíaco, as medidas de complexidade estão em foco pela comunidade científica, não só pela exigência existente nas características do sinal, mas principalmente pela capacidade de processamento e análise de dados que as tecnologias atuais providenciam e pela quantidade de dados existente.

O desafio que se apresenta prende-se na difícil interpretação das medidas não-lineares em contexto neurofisiológico, dificultando a integração de sistemas de apoio à decisão na prática clínica. Numa era em que os "wearables" começam a surgir, uma monitorização contínua irá proporcionar aos investigadores uma enorme quantidade de informação, proporcionando o desenvolvimento de modelos multivariados, como o "deep learning".

A Ciência da Complexidade tem uma grande importância na prevenção e previsão de desfechos de partos anormais, apoiando os profissionais de saúde nas suas decisões clínicas. Esta dissertação acrescentou um passo ao conhecimento atual de como a ciência da complexidade pode ser mais bem compreendida no contexto clínico e melhor utilizada na prevenção de maus desfechos neonatais.

Scientific Results

Published articles:

João Santos, Hernâni Gonçalves, João Bernardes, Luís Antunes, Mohammad Nozari, Cristina Santos. Entropy and Compression capture different complexity features, the case of fetal heart rate. *Entropy,* 2017, *19*(12), 688 <https://doi.org/10.3390/e19120688>. Published in the Special Issue Information Theory Applied to Physiological Signals

Impact Factor= 2.305 **ISI citations = 6**

João Santos, Teresa Henriques, Inês Nunes, Célia Costa, João Bernardes, Cristina Santos. Complexity of cardiotocographic signals as a predictor of labor. *Entropy* 2020, *22*(1), 104; <https://doi.org/10.3390/e22010104>. Published in the Special Issue Assessing Complexity in Physiological Systems through Biomedical Signals Analysis.

Impact Factor= 2.494 **ISI citations = 2**

Mohammad Nozari, Luísa Castro, João Santos, João Bernardes, Antónia Costa, Cristina Santos. On the prediction of fetal acidemia: a spectral analysis-based approach. Computers in Biology and Medicine, 2019, Volume 109, Pages 235-241. <https://doi.org/10.1016/j.compbiomed.2019.04.041>

Impact Factor= 3.434 **ISI citations = 3**

Prizes:

Best score in SPaM challenge organized by the Oxford University, in 2017

List of Figures

List of Tables

Acronyms

Acc	Number of accelerations
AIC	Akaike Information Criterion
ANS	Autonomic Nervous System
ApEn	Approximate Entropy
avLTV	Average Long-Term Variability
avSTV	Average Short-Term Variability
BE	Base excess deficit
CINTESIS	Centro de Investigação em Tecnologias e Serviços de Saúde, Faculdade de Medicina da Universidade do Porto
CRACS	Center for Research in Advanced Computing Systems, Faculdade de Ciências da Universidade do Porto
CTG	Cardiotocography
CR	Compression Ratio
CTU	Czech Technical University
CV	Coefficient of Variability
Dec	Number of Decelerations
DFT	Discrete Fourier Transform
ECG	Electrocardiogram
fECG	Fetal Electrocardiogram
FFT	Fast Fourier Transform
FHR	Fetal Heart rate
fHRV	Fetal Heart rate variability
FIGO	International Federation of Obstetrics and Gynaecology
HF	High Frequency
iDec	Number of intermediary Decelerations
IUGR	Intrauterine growth restriction
KC	Kolmogorov Complexity
LF	Low Frequency
LTI	Long Term Irregularity
LTV	Long Term Variability
LZ	Lempel Ziv
MA	Mildly Acidemic
mAge	Maternal Age
mDec	Number of mild Decelerations
MF	Movementfrequency
Medcids	Departamento de Medicina da Comunidade, Informação e Decisão em Saúde, Faculdade de Medicina da Universidade do Porto
MI	Mutual Information
MSA	Mild-to-Severly Acidemic
MSE	Multiscale Entropy
N	Normal
PSD	Power Spectral Density
ROC	Receiver operating characteristic
RMSSD	Root mean square of successive differences

SampEn	Sample Entropy
SD	Standard Deviation
SDNN	Standard deviation of normal-to-normal beat intervals
ShEn	Shannon Entropy
SPAM	Signal Processing and Monitoring in Labor
STV	Short-Term Variability
UAB	Umbilical artery blood
UHB	University Hospital of Brno
UC	Uterine contractions
VLF	Very Low frequency
%abLTV	Percentage of abnormal Long-Term variability
%abSTV	Percentage of abnormal Short-Term variability

INTRODUCTION

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 1/250 births in developed countries and 1/33 in South Asia and Sub-Saharan Africa (data from 2009) [1].

In developed countries, clinical decisions during labor are strongly based on fetal heart rate (FHR) monitoring [2, 3], being cardiotocography (CTG) [4] the most used tool to assess fetal well-being since the early '60s. CTG combines FHR measurement, obtained by means of a Doppler ultrasound probe and uterine contraction monitoring probe, recorded using an abdominal pressure transducer. The information provided by CTG is limited since a complete electrocardiogram (ECG) signal of the fetus is not available. Moreover, 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 ECG (fECG), containing not only FHR but also other relevant clinical parameters. On the other hand, fECG is only possible during labor after the beginning of cervical dilatation and rupture of the membranes, and therefore, carrying with it risks of infection [5, 6].

Other methods for fetal monitoring, such as fetal phonocardiography [7-9], fetal echocardiography [10, 11], and fetal magnetocardiography [12, 13], are used, and each one has its advantages and disadvantages. For more detail on this matter, see [14, 15].

The introduction of 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 show any significant improvement. Also, electronic fetal monitoring became the main suspect for the increased caesarean sections [16]. These procedures result in a slight increase in poor outcomes in low-risk pregnancies. They also require a longer time to heal than vaginal birth, and the increased risks include baby breathing problems, amniotic fluid embolism, and postpartum bleeding for the mother [17].

Despite the importance of the fetus and mother's well-being assessment, low concordance between physicians is still present, even among experienced obstetricians, resulting in a high

false-positive rate [2, 18, 19]. As in daily practice, FHR is visually interpreted by the clinician, even when following the guidelines provided by the International Federation of Obstetrics and Gynecology (FIGO) [20, 21], it is associated with high sensitivity but low specificity [22]. The low specificity might lead to a chance of more harmful than beneficial adherence to conventional guidelines as it creates unnecessary interventions [23]. This issue may be related to Plesk and Greenhalgh's conjecture, which states that healthcare components belong to the simple, complex, or chaotic domains. In particular, the analysis of fetal heart rate variability (fHRV) is somewhat between the complex and chaotic domains, reflecting the low certainty and agreement present in the literature (as shown in the certainty-agreement diagram in Figure 1).

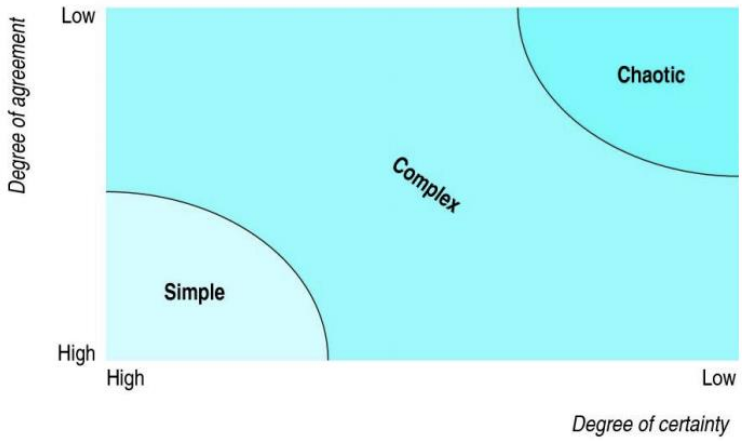

Figure 1.1. Certainty-agreement diagram

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 [6]. A certain level of unpredictable fHRV reflects sufficient capabilities of the organism in search of optimal behavior. Reduced fHRV is linked with limited capabilities and mental disorders [24]. Some linear modeling approaches quantify sympathetic and parasympathetic control mechanisms and their balance through 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 [25]. 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 nonlinear parameters [26, 27]. This approach can reveal relevant clinical information not exposed by temporal or frequency analysis [28].

The FHR variability analysis and the spectral analysis and nonlinear methods used in this dissertation will be described in detail in the following sections.

Fetal Heart Rate Variability Analysis

The purpose of fetal monitoring resides in assuring normal fetal development and the timely identification of fetal and maternal well-being compromise. It is crucial to prevent intrapartum fetal hypoxia, 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 time. This situation leads to an accumulation of carbon dioxide leading to fetal acidemia, resulting in a lower pH in the fetal blood vessels. The early detection of babies at risk of acidemia might decrease the chance of a post-diagnosis of cerebral palsy, neonatal encephalopathy, or even death [29]. As state before, conventional analysis of CTG provides information about cortical and ANS activity [6]. However, CTG has poor specificity, leading to a high rate of false positives. This leads to unnecessary operative interventions, increased caesarean section rate, without any clear benefit to the perinatal outcome, disrupting natural procedures that imply certain risks for both mother and baby. Under this spectrum, additional tests such as fetal scalp blood sampling, fetal pulse oximetry, and fetal electrocardiograph (fECG, also called STAN or ST-analyzer) can be performed [30]. In particular, the fECG is a graphic record of the myocardial cells' electrical activity, reflecting the myocardium's oxygenation level.

STAN assesses changes in the fECG complex, particularly in the ST segment and T wave (Figure 2.2). They are related to electrical changes that occur when the next myocardium contraction is being prepared. Besides, the T/QRS ratio is also analyzed, reflecting the duration of the hypoxic insult in the cells [31]. These inputs give information to calculate physiological

features such as the baseline, which, in return, defines accelerations and decelerations. Systems such as Omniview SisPorto [32], OxSys [33], and CAFE [34] automatically deal with CTG assessment.

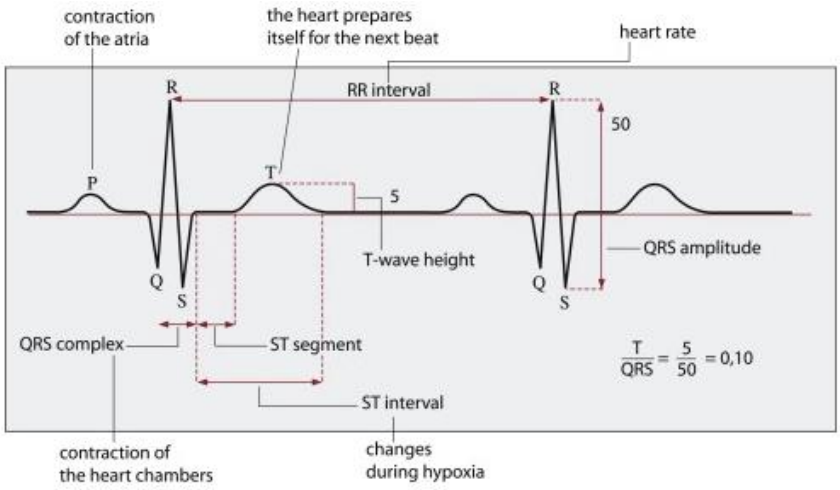

Figure 1.2. Fetal Electrocardiogram Complex [31]

Linear features usually analyzed comprise: FHR baseline, that is the mean FHR in the absence of uterine contractions and fetal movement, and it retrieves information mainly on the intrinsic activity of the heart [35]; the interaction between sympathetic and parasympathetic systems, which express the influence of external and internal stimuli, are captured by short (STV) and long term variability (LTV) [6], respectively, and they are defined as the beat-to-beat changes over one minute, and the variation in interval length over a certain number of RR intervals (time difference between two consecutive R peaks of the fECG complex); accelerations are the increase in the fetal heart rate over the fetal baseline heart rate and reflect the sympathetic activity, while decelerations are the decrease in the fetal heart rate below the fetal baseline heart rate and are conveyed by the parasympathetic system [6]; the standard deviation of normal-to-normal beat intervals (SDNN) reflects sympathetic and vagal heart rate modulations [36]; skewness, quantifying the symmetry of a series; kurtosis, an outlier quantification measure; root mean square of successive differences (RMSSD) reflects vagal control [37]; the ratio SDNN/RMSSD, is a temporal alternative for the frequency LF/HF

measure (ration between Low Frequency and High Frequency) as a marker of the sympatho-vagal balance [38]. Other features in the time domain are also widely used, such as the long-term irregularity (LTI), the Interval index, and the delta value of the FHR signal [39, 40].

Spectral Analysis

At the beginning of the 80s, Akselrod et al. [41] introduced the power spectral density (PSD) analysis to assess beat-to-beat cardiovascular control quantitatively. Frequency domain analysis was relevant for understanding the autonomic background of RR interval fluctuations [42, 43]. PSD methods are usually classified as parametric or nonparametric. Although both strategies achieve similar results in most cases, nonparametric methods (Fast Fourier Transform (FFT) for the majority of the cases) have the advantage of simplicity and higher processing speed. On the other hand, parametric methods achieve a better estimation of PSD, even for a small number of samples. The concern with parametric methods is the need to guarantee the signal's stationarity, which does not happen in the case of fHR [36].

The spectral analysis started with the work of the mathematician Fourier (1807), who suggested that any periodic function could be estimated using combinations of sinus and co-sinus functions.

For a given function f(x) (for instance, hear rate series over time x), its Fourier series is:

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

where the coefficients are calculated as followed:

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x)\, dx$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \;\; dx$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \;\; dx$$

This notion was the first step for a time series analysis in the frequency spectrum. In the case of fHR, where data is discrete and equally spaced (as is for time intervals), the Discrete Fourier Transform (DFT) takes place. A given equally spaced sequence $x_0 \dots x_{n-1}$, of size n is transformed, from a time domain to a frequency domain, in a sequence of complex numbers:

$$X_k = \sum_{j=0}^{n-1} x_j e^{-\frac{2\pi i k j}{n}},$$

with $i = \sqrt{-1}$.

Applying DFT comes with a cost of n^2 operations, the procedure becomes costly as the sample size rises. The FFT achieves better results as the cost decreases to $n \log n$ operations [44]. The most used algorithm to apply FFT is the Cooley-Tukey algorithm [45], a divide and conquer algorithm that recursively decomposes a DFT into smaller ones, along with linear multiplications of i.

The usual spectral analysis considers the power of the signal over the spectrum of frequencies. Power spread over all frequencies means high irregularity, while distinct peaks over specific frequencies indicate predominant rhythms and regular patterns [46]. Frequently it consists of gathering data from three main components: Very Low Frequency (VLF, $\leq 0.04Hz$), Low Frequency (LF, $0.04 - 0.15Hz$), and High Frequency (HF, $0.15 - 0.4Hz$). The high-frequency component is mainly indicating parasympathetic nervous system activity. The sympathetic nervous system activity is reflected by the lower frequency components (however, traces of parasympathetic activity can also be found at lower frequency levels). Therefore, the ratio of energies $\frac{LF}{HF}$ is also widely used. It expresses the balance of behavior of sympathetic and parasympathetic branches of the ANS in adults for short-term recordings [36].

An alternative branching was proposed by Signorini [47], where the new bin Movement Frequency (MF, $0.15 - 0.5Hz$), related to fetal movement and maternal breathing, was defined. From this approach, VLF ($\leq 0.03Hz$) was related to long periods of nonlinear contributions, LF ($0.03 - 0.15Hz$) was associated with the neural sympathetic activity, and

HF ($0.5-1\ Hz$) correlated with fetal breathing. Here, the ratio of energies was defined as $\frac{LF}{MF+HF}$.

Reinhard [48] studied the HRV changes during hypnosis and noted increased LF power and a decreased HF power, meaning a sympathovagal shift towards increased sympathetic modulation, compared to pregnant women at rest.

Spectral analysis has also been important in the study of maturation. $\frac{LF}{HF}$ was significantly reduced in intrauterine growth restriction (IUGR) fetuses compared with a control group [49].

Nonlinear Methods

Nonlinear phenomena are intrinsic to HRV by the complex interactions of electrophysiological, hemodynamic, and humoral agents, particularly the autonomic and central nervous system regulations.

Variability and complexity are distinct terms. While a complex system requires variability, the other way around is not guaranteed. For example, a set of random notes in a piece of music can be interpreted as having high complexity, for its non-predictability. In contrast, 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 and/or spatial structures over a varied range of scales. [50]. In the end, complexity is a property of any system that quantifies the amount of structured information.

From this perspective, numerous approaches have been suggested in the past decades. Although this work's focus is only on entropy and compression, other nonlinear methods such as fractal analysis and wavelets have been the object of studies when applied to fHR.

Entropy

According to Shannon [51], 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 series's irregularity, has been widely used in physiological signal analysis.

The focus will be on Shannon Entropy (ShEn), since it is the root of all others, and Approximate Entropy (ApEn), Sample Entropy (SampEn), and Multiscale Entropy (MSE), as they were used in the papers.

Shannon entropy (ShEn)

Though ShEn was introduced in 1948, some authors still applied it to analyze FHRV [52, 53]. Considering a time series $X = \{x_i\}$, of N points, ShEn $H(X)$ is a functional of its joint probability density function $p(X)$ defined as [51]:

$$H(X) = -\sum_{i}^{n} p\big(x(i)\big) \ln(p\big(x(i)\big))$$

Where $p(x(i))$ is the probability of $X = x(i)$. Shannon entropy measures the total information contained in the process X.

Approximate entropy (ApEn)

In 1991, Pincus developed a regularity statistic tool to quantify a system's complexity, based on entropy [54, 55]. The measure is based on the theory that healthy dynamic stability comes from specific mechanisms and properties of interconnected networks. When a weak connection arises between systems or within one, it is the disease mechanism, which is characterized by an increase of regularity of the series [56].

ApEn is defined as the logarithmic likelihood that the series's patterns closer to each other will remain close when the next comparison with a longer pattern is made. It measures the irregularity of a time series robustly with short segments [57] as follows:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r)$$

$$= \frac{1}{N-m+1}\sum_{i=1}^{N-m+1} \ln\left(C_r^m(i)\right) - \frac{1}{N-m}\sum_{i=1}^{N-m} \ln(C_r^{m+1}(i)),$$

$$C_r^m = \frac{n_i^m(r)}{N-m}$$

where $n_i^m(r)$ is the number of patterns in p_m (all patterns of length m) that are similar to $p_m(i)$ (subsequence of m consecutive signal values, beginning in instance i). C_r^mrepresents the correlation integral, m is the embedding dimension and tolerance r working as a threshold, for N signal points.

Two patterns are considered similar if the difference between any pair of corresponding measurements is less than or equal to r. Values of 0.1, 0.15, or 0.2 standard deviations are usually used for parameter r, while m is mostly considered as 2 [58].

Sample entropy (SampEn)

SampEn was proposed by Richman and Moorman (2000) [59, 60] in a similar manner as ApEn, but with some major differences, reducing bias, especially in short data sets as it is dependent on the record length. Besides, SampEn eliminates self-matches, and the conditional probabilities are not estimated in a template manner.

This probability measure is computed directly as the logarithm of conditional probability and not from the logarithmic sums ratio. Considering the same m and r parameters from ApEn, A as the number of pairs of vectors X with length $m+1$, with $d[X_m(i), X_m(j)] \leq r, i \neq j$, and B as the number of template matches of the size m, SampEn is defined as:

$$SampEn = -ln\frac{A}{B}$$

Multiscale entropy (MSE)

ApEn and SampEn have the disadvantage of outputting a single index concerning the series's general behavior, thus not revealing its underlying dynamics. MSE is able to provide this information since several indices are given, one for each scale in the time series.

Considering a time series $\{x_i\}$ of N points, it constructs consecutive coarse-grained time series $\{y^{(\tau)}\}$, as a function of the factor τ,

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i\,, 1 \leq j \leq \frac{N}{\tau}.$$

$y^{(1)}$ is the original time series, and $\frac{N}{\tau}$ the length of each coarse-grained series [61].

Some authors support using a tolerance r obtained from the original series and keep it constant for all scales [62]. Other authors favor choosing an individual tolerance level r for each scale [63, 64]. For example, the quadratic sample entropy permits a personalized estimation of r for each scale in short data [65].

MSE has been widely employed in biomedical signal analysis as it allows measuring signal properties at different time scales [66].

The physiological interpretation of multiscale complexity is not very clear because, in a complex dynamic system, all scales might be affected by regulating influences [67]. Low complexity scales indicate regular patterns with periodicity, but isolated ones would indicate the periodicity of one single frequency oscillation and that usually is not present in complex systems. However, it is typical of the appearance of correlated neighboring scales [62, 68].

Other entropies

The literature also contains usage of different entropies and some are still emerging. Examples include entropy of first order Markov model [52], transfer entropy [69], Kullback-Leibler entropy [70], mutual information (MI) [67, 71-73], Rényi entropy [53, 74-76], Kernel-based entropy [75] Bubble entropy [76], S/k ratio of entropy [77], δ-entropy [78] and permutation entropy [38, 68, 79, 80].

Compression

Dynamic systems theory was firstly linked with information theory by Kolmogorov [81] in 1958. Years later, "algorithm information theory" was then independently proposed by three different authors (Solomonoff [82], Kolmogorov [83] and Chaitin [84]).

The Kolmogorov Complexity (KC) K is defined as the function mapping a string x in an integer, bounded to a Turing Machine ϕ:

$$K_\phi(x) = \begin{cases} min\{|p|: \phi(p) = x\}, & if\ \phi(p) = x \\ \infty & if\ p\ does\ not\ exit \end{cases}$$

The KC reflects the increase of new patterns along a given sequence. In this case, the word complexity refers to the algorithmic complexity, defined according to Information Theory, as the shortest program's length p able to print the string x. 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 function will be the original string's length as any sort of 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 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.

Compressors could be divided into two big groups: lossless or lossy. The former group is composed of compressors, in which every bit of data after decompressed is restored. For the lossy group, this is not guaranteed, particularly for redundant information.

Innumerous compressors are found in the literature. Lempel-Ziv (LZ) [85] was introduced by Lempel and Ziv in 1976 and is the starting point for different compressors, such as the LZ77, LZ78, and gzip. bzip2 was developed by Julian Seward and uses the block sort algorithm giving very fast results; PPM stands for Prediction by Partial Matching and belongs to a set of more recent compressors using statistic models. However, it is slow and computationally

demanding. Compressors PAQ also falls short for these characteristics, achieving very high compressibility levels using neural networks.

In order to estimate the complexity of a physiological signal using compression, different approaches have been used, such as the heartbeat series [86], an increase/decrease coding system using a binary [28, 87, 88], or ternary [89, 90] alphabet.

The applications of compression in health research range from event detection (such as epileptic seizure [91], the onset of ventricular tachycardia or fibrillation [92] and changes from sleep to waking state in-depth anesthesia [93]), characterizing neural spike trains [94] or in DNA sequences studies [95].

Objectives

The primary purpose of this thesis is to explore how complexity analysis relates to FHRV and how it can be used to prevent bad neonatal outcomes, as pre-term birth and fetal asphyxia. For this matter, more specific goals were defined:

- How do complexity measures relate to standard physiological characteristics of FRH tracings and how correlated are they with each other?
- How well can complexity measures predict labor?
- How well can complexity measures predict fetal acidemia?
- Is it possible to develop new methods for labor and fetal acidemia prediction?

References

1. Cousens, S., et al., National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. Lancet, 2011. **377**(9774): p. 1319-30.

2. Ayres-De-Campos, D., et al., Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. British Journal of Obstetrics and Gynaecology, 1999. **106**(12): p. 1307-1310.

3. Bernardes, J. and D. Ayres-de-Campos, The persistent challenge of foetal heart rate monitoring. Current Opinion in Obstetrics & Gynecology, 2010. **22**(2): p. 104-109.

4. Spencer, J.A., Role of cardiotocography. Br J Hosp Med, 1992. **48**(2): p. 115-8.

5. Bartlett, M.L., A. Murray, and W. Dunlop, Is fetal heart rate monitoring sufficiently sensitive to detect changes during labour? J Biomed Eng, 1992. **14**(5): p. 431-4.

6. Parer, J.T., Handbook of Fetal Heart Rate Monitoring. 1997: W.B. Saunders Company.

7. Kovacs, F., et al., Fetal phonocardiography--past and future possibilities. Comput Methods Programs Biomed, 2011. **104**(1): p. 19-25.

8. Adithya, P.C., et al., Trends in fetal monitoring through phonocardiography: Challenges and future directions. Biomedical Signal Processing and Control, 2017. **33**: p. 289-305.

9. Chourasia, V.S.T., Anil Kumar, FETAL HEART RATE VARIABILITY ANALYSIS FROM PHONOCARDIOGRAPHIC RECORDINGS. 2011.

10. Persico, N., et al., Fetal echocardiography at 11-13 weeks by transabdominal high-frequency ultrasound. Ultrasound Obstet Gynecol, 2011. **37**(3): p. 296-301.

11. Zhang, D.Y., 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-a Journal of Cardiovascular Ultrasound and Allied Techniques, 2016. **33**(1): p. 90-98.

12. Quartero, H.W., et al., Clinical implications of fetal magnetocardiography. Ultrasound Obstet Gynecol, 2002. **20**(2): p. 142-53.

13. Grimm, B., et al., Recommended standards for fetal magnetocardiography. Pacing Clin Electrophysiol, 2003. **26**(11): p. 2121-6.

14. Jaros, R., R. Martinek, and R. Kahankova, Non-Adaptive Methods for Fetal ECG Signal Processing: A Review and Appraisal. Sensors, 2018. **18**(11).

15. Hoyer, D., et al., Monitoring fetal maturation-objectives, techniques and indices of autonomic function. Physiol Meas, 2017. **38**(5): p. R61-r88.

16. Steer, P.J., Has electronic fetal heart rate monitoring made a difference? Seminars in Fetal & Neonatal Medicine, 2008. **13**(1): p. 2-7.

17. Caughey, A.B., et al., Safe prevention of the primary cesarean delivery. American Journal of Obstetrics and Gynecology, 2014. **210**(3): p. 179-193.

18. Bernardes, J., et al., Evaluation of interobserver agreement of cardiotocograms. International Journal of Gynecology & Obstetrics, 1997. **57**(1): p. 33-37.

19. Donker, D.K., H.P. Vangeijn, and A. Hasman, Interobserver Variation in the Assessment of Fetal Heart-Rate Recordings. European Journal of Obstetrics & Gynecology and Reproductive Biology, 1993. **52**(1): p. 21-28.

20. Huch, A., R. Huch, and G. Rooth, 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. European Journal of Obstetrics Gynecology and Reproductive Biology, 1994. **54**(3): p. 165-175.

21. Ayres-De-Campos, D., et al., FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. International Journal of Gynecology & Obstetrics, 2015. **131**(1): p. 13-24.

22. Schiermeier, S., et al., Sensitivity and specificity of intrapartum computerised FIGO criteria for cardiotocography and fetal scalp pH during labour: multicentre, observational study. BJOG, 2008. **115**(12): p. 1557-63.

23. Plsek, P.E. and T. Greenhalgh, Complexity science - The challenge of complexity in health care. British Medical Journal, 2001. **323**(7313): p. 625-628.

24. Rotmensch, S., et al., The effect of betamethasone and dexamethasone on fetal heart rate patterns and biophysical activities - A prospective randomized trial. Acta Obstetricia Et Gynecologica Scandinavica, 1999. **78**(6): p. 493-500.

25. Signorini, M.G., A. Fanelli, and G. Magenes, Monitoring fetal heart rate during pregnancy: contributions from advanced signal processing and wearable technology. Comput Math Methods Med, 2014. **2014**: p. 707581.

26. Nunes, I., et al., An overview of central fetal monitoring systems in labour. Journal of Perinatal Medicine, 2013. **41**(1): p. 93-99.

27. Wilson, T. and T. Holt, Complexity science - Complexity and clinical care. British Medical Journal, 2001. **323**(7314): p. 685-688.

28. Chudacek, V.S., Jiri; Huptych, Michal; Georgoulas, George; Janku, Petr; Koucky, Michal; Stylios, Chrysostomos; Lhotska, Lenka, Automatic Classification of Intrapartal Fetal Heart-Rate Recordings - Can It Compete with Experts? 2010.

29. Georgoulas, G., C. Stylios, and P. Groumpos, Feature extraction and classification of Fetal Heart Rate using wavelet analysis and Support Vector Machines. International Journal on Artificial Intelligence Tools, 2006. **15**(3): p. 411-432.

30. East, C.E., et al., Fetal pulse oximetry for fetal assessment in labour. Cochrane Database of Systematic Reviews, 2014(10).

31. Amer-Wahlin, I. and K. Marsal, ST analysis of fetal electrocardiography in labor. Seminars in Fetal & Neonatal Medicine, 2011. **16**(1): p. 29-35.

32. Ayres-de-Campos, D., et al., SisPorto 4.0-computer analysis following the 2015 FIGO Guidelines for intrapartum fetal monitoring. Journal of Maternal-Fetal & Neonatal Medicine, 2017. **30**(1): p. 62-67.

33. Georgieva, A., et al., Artificial neural networks applied to fetal monitoring in labour. Neural Computing & Applications, 2013. **22**(1): p. 85-93.

34. Guijarro-Berdinas, B., A. Alonso-Betanzos, and O. Fontenla-Romero, Intelligent analysis and pattern recognition in cardiotocographic signals using a tightly coupled hybrid system. Artificial Intelligence, 2002. **136**(1): p. 1-27.

35. Bernardes, J., et al., A more objective fetal heart rate baseline estimation. Br J Obstet Gynaecol, 1996. **103**(7): p. 714-5.

36. Camm, A.J., et al., Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. European Heart Journal, 1996. **17**(3): p. 354-381.

37. Moraes, E.R., et al., Linear and nonlinear measures of fetal heart rate patterns evaluated on very short fetal magnetocardiograms. Physiol Meas, 2012. **33**(10): p. 1563-83.

38. Schneider, U., et al., Human fetal heart rate variability-characteristics of autonomic regulation in the third trimester of gestation. J Perinat Med, 2008. **36**(5): p. 433-41.

39. Chudacek, V., et al., Assessment of features for automatic CTG analysis based on expert annotation. Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 6051-4.

40. Giuliano, N., et al., Computerised analysis of antepartum foetal heart parameters: New reference ranges. J Obstet Gynaecol, 2017. **37**(3): p. 296-304.

41. Akselrod, S., et al., Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. Science, 1981. **213**(4504): p. 220-2.

42. Pomeranz, B., et al., Assessment of autonomic function in humans by heart rate spectral analysis. Am J Physiol, 1985. **248**(1 Pt 2): p. H151-3.

43. Pagani, M., et al., Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. Circ Res, 1986. **59**(2): p. 178-93.

44. Tinku Acharya, A.k.R., Image Processing Principles and Applications. 2005: John Wiley.

45. Cooley, J.W. and J.W. Tukey, On the Origin and Publication of the Fft Paper - a Citation-Classic Commentary on an Algorithm for the Machine Calculation of Complex Fourier-Series - Cooley,J.W., Tukey,J.W. Current Contents/Engineering Technology & Applied Sciences, 1993(51-52): p. 8-9.

46. Hoyer, D., et al., Fetal development of complex autonomic control evaluated from multiscale heart rate patterns. Am J Physiol Regul Integr Comp Physiol, 2013. **304**(5): p. R383-92.

47. Signorini, M.G., et al., Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings. IEEE Trans Biomed Eng, 2003. **50**(3): p. 365-74.

48. Reinhard, J., et al., Change of Spectral Analysis of Fetal Heart Rate During Clinical Hypnosis: a Prospective Randomised Trial from the 20th Week of Gestation Till Term. Geburtshilfe Und Frauenheilkunde, 2012. **72**(4): p. 316-321.

49. Vinkesteijn, A.S.M., et al., Fetal heart rate and umbilical artery flow velocity variability in intrauterine growth restriction: a matched controlled study. Ultrasound in Obstetrics & Gynecology, 2004. **23**(5): p. 461-465.

50. Schnettler, W.T., et al., Complexity analysis of fetal heart rate preceding intrauterine demise. Eur J Obstet Gynecol Reprod Biol, 2016. **203**: p. 286-90.

51. Shannon, C.E., A Mathematical Theory of Communication. Bell System Technical Journal, 1948. **27**(3): p. 379-423.

52. Hopkins, P., et al., A comparative study of fetal heart rate variability analysis techniques. Conf Proc IEEE Eng Med Biol Soc, 2006. **1**: p. 1784-7.

53. Khandoker, A.H., et al., Analysis of maternal-fetal heart rate coupling directions with partial directed coherence. Biomedical Signal Processing and Control, 2016. **30**: p. 25-30.

54. Pincus, S.M., Approximate entropy as a measure of system complexity. Proc Natl Acad Sci U S A, 1991. **88**(6): p. 2297-301.

55. Pincus, S.M. and R.R. Viscarello, Approximate entropy: a regularity measure for fetal heart rate analysis. Obstet Gynecol, 1992. **79**(2): p. 249-55.

56. Buchman, T.G., The community of the self. Nature, 2002. **420**(6912): p. 246-251.

57. Fusheng, Y.a.B., H and Qingyu, T, Approximate entropy and its application in biosignal analysis. Nonlinear Biomedical Signal Processing: Dynamic Analysis and Modeling, 2001. **2**: p. 72-91.

58. Goncalves, H., et al., Frequency domain and entropy analysis of fetal heart rate: appealing tools for fetal surveillance and pharmacodynamic assessment of drugs. Cardiovasc Hematol Disord Drug Targets, 2008. **8**(2): p. 91-8.

59. Richman, J.S. and J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology-Heart and Circulatory Physiology, 2000. **278**(6): p. H2039-H2049.

60. Lake, D.E., et al., Sample entropy analysis of neonatal heart rate variability. American Journal of Physiology-Regulatory Integrative and Comparative Physiology, 2002. **283**(3): p. R789-R797.

61. Costa, M., A.L. Goldberger, and C.K. Peng, Multiscale entropy to distinguish physiologic and synthetic RR time series. Computers in Cardiology 2002, Vol 29, 2002. **29**: p. 137-140.

62. Costa, M., A.L. Goldberger, and C.K. Peng, Multiscale entropy analysis of biological signals. Physical Review E, 2005. **71**(2).

63. Valencia, J.F., et al., Refined Multiscale Entropy: Application to 24-h Holter Recordings of Heart Period Variability in Healthy and Aortic Stenosis Subjects. Ieee Transactions on Biomedical Engineering, 2009. **56**(9): p. 2202-2213.

64. Nikulin, V.V. and T. Brismar, Comment on "Multiscale entropy analysis of complex physiologic time series". Physical Review Letters, 2004. **92**(8).

65. Lake, D.E. and J.R. Moorman, Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. American Journal of Physiology-Heart and Circulatory Physiology, 2011. **300**(1): p. H319-H325.

66. Signorini, M.G., et al., Advances in monitoring cardiovascular signals. Contribution of nonlinear signal processing. Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 6568-71.

67. Gieraltowski, J., et al., Development of multiscale complexity and multifractality of fetal heart rate variability. Auton Neurosci, 2013. **178**(1-2): p. 29-36.

68. Hoyer, D., et al., Fetal development assessed by heart rate patterns--time scales of complex autonomic control. Comput Biol Med, 2012. **42**(3): p. 335-41.

69. Marzbanrad, F., et al., Transfer entropy analysis of maternal and fetal heart rate coupling. Conf Proc IEEE Eng Med Biol Soc, 2015. **2015**: p. 7865-8.

70. Tetschke, F., et al., Assessment of fetal maturation age by heart rate variability measures using random forest methodology. Comput Biol Med, 2016. **70**: p. 157-62.

71. Hoyer, D., S. Nowack, and U. Schneider, Multi-scale characteristics of resampled fetal heart rate pattern provide novel fetal developmental indices. Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 1451-4.

72. Granero-Belinchon, C., et al., Information Theory to Probe Intrapartum Fetal Heart Rate Dynamics. Entropy, 2017. **19**(12).

73. Spyridou, K., et al., 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 Meas, 2018. **39**(1): p. 015007.

74. Lake, D.E., Renyi entropy measures of heart rate Gaussianity. Ieee Transactions on Biomedical Engineering, 2006. **53**(1): p. 21-27.

75. Zaylaa, A.J., et al., Cascade of Nonlinear Entropy and Statistics to Discriminate Fetal Heart Rates. 2016 3rd International Conference on Advances in Computational Tools for Engineering Applications (Actea), 2016: p. 152-157.

76. Manis, G. and R. Sassi, Relation between fetal HRV and value of umbilical cord artery pH in labor, a study with entropy measures. 2017 Ieee 30th International Symposium on Computer-Based Medical Systems (Cbms), 2017: p. 272-277.

77. Velasquez, J.O.R., M.F.C. Zamora, and R.C. Marin, Mathematical diagnosis of fetal monitoring tracing based on the S/k ratio of entropy. Archivos De Medicina, 2017. **17**(1): p. 27-37.

78. Oudjemia, S., et al., Delta-Fuzzy Similarity Entropy to Discriminate Healthy from Sick Fetus. 2013 2nd International Conference on Advances in Biomedical Engineering (Abme 2013), 2013: p. 1-4.

79. Frank, B., et al., Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses. Medical & Biological Engineering & Computing, 2006. **44**(3): p. 179-187.

80. Barquero-Perez, O., et al., Fetal Heart Rate Complexity Measures to Detect Hypoxia. 2015 Computing in Cardiology Conference (Cinc), 2015. **42**: p. 133-136.

81. Kolmogorov, A.N., A New Metric Invariant of Transient Dynamical Systems and Automorphisms in Lebesgue Spaces. Doklady Akademii Nauk Sssr, 1958. **119**(5): p. 861-864.

82. Solomonoff, R.J., Formal Theory of Inductive Inference .2. Information and Control, 1964. **7**(2): p. 224-+.

83. Kolmogorov, A.N., Three approaches to the definition of the concept “quantity of information. Probl. Peredachi Inf., 1965.

84. Chaitin, G.J., On Length of Programs for Computing Finite Binary Sequences - Statistical Considerations. Journal of the Acm, 1969. **16**(1): p. 145-&.

85. Lempel, A. and J. Ziv, On the Complexity of Finite Sequences. IEEE Transactions on Information Theory, 1976. **22**(1): p. 75-81.

86. Monteiro-Santos, J., et al., Entropy and Compression Capture Different Complexity Features: The Case of Fetal Heart Rate. Entropy, 2017. **19**(12).

87. Yang, A.C.C., et al., Linguistic analysis of the human heartbeat using frequency and rank order statistics. Physical Review Letters, 2003. **90**(10).

88. Spilka, J., et al., Using nonlinear features for fetal heart rate classification. Biomedical Signal Processing and Control, 2012. **7**(4): p. 350-357.

89. Ferrario, M., et al., Multiparameter analysis of heart rate variability signal for the investigation of high risk fetuses. Conf Proc IEEE Eng Med Biol Soc, 2009. **2009**: p. 4662-5.

90. Magenes, G., et al., Multivariate analysis based on linear and non-linear FHR parameters for the identification of IUGR fetuses. Conf Proc IEEE Eng Med Biol Soc, 2014. **2014**: p. 1868-71.

91. Radhakrishnan, N. and B.N. Gangadhar, Estimating regularity in epileptic seizure time-series data. Ieee Engineering in Medicine and Biology Magazine, 1998. **17**(3): p. 89-94.

92. Zhang, X.S., et al., Detecting ventricular tachycardia and fibrillation by complexity measure. Ieee Transactions on Biomedical Engineering, 1999. **46**(5): p. 548-555.

93. Zhang, X.S., R.J. Roy, and E.W. Jensen, EEG complexity as a measure of depth of anesthesia for patients. Ieee Transactions on Biomedical Engineering, 2001. **48**(12): p. 1424-1433.

94. Szczepanski, J., et al., Application of Lempel-Ziv complexity to the analysis of neural discharges. Network-Computation in Neural Systems, 2003. **14**(2): p. 335-350.

95. Orlov, Y.L. and V.N. Potapov, Complexity: an internet resource for analysis of DNA sequence complexity. Nucleic Acids Research, 2004. **32**: p. W628-W633.

OUTLINE

2. Outline

After the Introduction (**Chapter 1**) presented above, and the Outline presented here (**Chapter 2**), the rest of the thesis is organized in four chapters that describe the work performed, three of which presenting the three studies already published in the scientific literature and a last chapter presents a general discussion and conclusions. A description of these chapters follows.

Chapter 3 aims to evaluate the association between nonlinear measures, entropy and compression, and the linear measures used by the Omniview-SisPorto program, an automatic fetal heart rate analysis software. This chapter also analyses parameter changes and how they behave at different scales (multi-scale analysis). It is also evaluated how these measures distinguish fetuses that at birth were considered acidemic or not, based on the pH of the blood collected from the umbilical cord.

In **Chapter 4,** intends to assess how nonlinear measures can predict whether or not the fetus is close to the time of delivery (with forecasts at one and two weeks). The issue is relevant because of the clinical importance of identifying cases where the risk of prematurity.

In order to explore other ways of predicting fetal asphyxia and following the participation of the working group in Signal Processing and Monitoring (SPaM) 2017, which consisted of the development of an algorithm for the prediction of acidemia using fetal heartbeat obtained during childbirth, a model based on spectral analysis was developed and subsequently validated in two different data sets (**Chapter 5**).

The dissertation ends in **Chapter 6** with an overall discussion and main conclusions.

ENTROPY AND COMPRESSION CAPTURE DIFFERENT COMPLEXITY FEATURES: THE CASE OF FETAL HEART RATE

João Santos, Cristina Santos, Luís Antunes, João Bernardes, Mohammad Nozari, Hernâni Gonçalves

ENTROPY

SPECIAL ISSUE "INFORMATION THEORY APLIED TO PHYSIOLOGICAL SIGNALS"

DECEMBER 2017

3. Entropy and compression capture different complexity features: the case of fetal heart rate

Abstract

Entropy and compression have been used to distinguish fetuses at risk of hypoxia from their healthy counterparts through the analysis of Fetal Heart Rate (FHR). Low correlation that was observed between these two approaches suggests that they capture different complexity features. This study aims at characterizing the complexity of FHR features captured by entropy and compression, using as reference international guidelines. Single and multi-scale approaches were considered in the computation of entropy and compression. The following physiologic-based features were considered: FHR baseline; percentage of abnormal long (%abLTV) and short (%abSTV) term variability; average short-term variability; and, number of acceleration and decelerations. All of the features were computed on a set of 68 intrapartum FHR tracings, divided as normal, mildly, and moderately-severely acidemic born fetuses. The correlation between entropy/compression features and the physiologic-based features was assessed. There were correlations between compressions and accelerations and decelerations, but neither accelerations nor decelerations were significantly correlated with entropies. The %abSTV was significantly correlated with entropies (ranging between −0.54 and −0.62), and to a higher extent with compression (ranging between −0.80 and −0.94). Distinction between groups was clearer in the lower scales using entropy and in the higher scales using compression. Entropy and compression are complementary complexity measures.

Introduction

In developed countries, clinical decisions during labor are strongly based on Fetal Heart Rate (FHR) monitoring [1,2], and cardiotocography is the tool that is routinely used for FHR and uterine contractions recordings. FHR is generally assessed in beats per minute to evaluate

fetal well-being allowing for an obstetrician to intervene and prevent potentially irreversible fetal brain damage or death. Despite the importance of FHR monitoring, poor reproducibility of visual analysis of cardiotocograms have been reported [1,3], and consequently computerized FHR analysis and new signal processing and pattern recognition techniques have been developed [4–6]. In this setting, complexity analysis of FHR recordings remains one of the most challenging tasks. Actually, FHR during labor seems to be part of a complex system, where, most of the times, individual agents behave in unpredictable ways, and whose actions are connected, inducing changes to one another [7]. In cases like this, a high degree of uncertainty is known to be present, leading to a poor interrater agreement. As a result, uncritical adherence to conventional guidelines might become more harmful than beneficial [7] and other approaches may be more appropriate, such as nonlinear models and scan of patterns [8].

Complexity is a property of systems that quantifies the amount of structured information and may be assessed using both entropy and compression. Approximate entropy (ApEn) is a measure of complexity, introduced by Pincus, used to quantify the amount of regularity and the unpredictability of fluctuations over time-series [9]. Later, Sample Entropy (SampEn) was presented by Richman and Moorman with the same goal as ApEn to assess biological time series [10]. In the particular case of FHR analysis, ApEn and SampEn are the most used measures of complexity, and are known to be used in the detection of different pathologies. On the other hand, the Kolmogorov complexity of an object is the length of the shortest computer program that can output it. Although Kolmogorov complexity is a non-computable measure, compressors do a very good job approximating it. This approach has led to positive results in very different subjects, as in literature [11], music [12], and computer virus and Internet traffic analysis [13]. Despite the successful application of compressors to FHR tracings in pathology detection, they have been used only to a limited extent in the analysis of biological signals to date [6,14].

Although both entropy and compression were able to distinguish fetuses at risk of hypoxia from their healthy counterparts through the analysis of the FHR signal, their low correlation suggests that these measures capture different features [15]. Henrique’s work also suggested

further research in order to study how physiological features are captured by entropy and compression.

The small computational time that is associated with both measures, namely with compressors, is particularly convenient if their inclusion in existing FHR monitoring systems is justified. Hopefully, the information on the fetus complexity obtained from the FHR signal may provide important auxiliary information to clinicians in fetal assessment, supporting clinical decisions. However, as entropy and compression seem to capture different features, it is important to study which features are captured by these measures. Following previous work's suggestion [15], this study aims at characterizing the complexity FHR features that are captured by entropy and by compression having as reference international clinical guidelines, and exploring the multiscale approach for entropy and compression.

Methods

Sixty eight FHR intrapartum tracings consecutively selected from a pre-existing database of term singleton gestations, with at most 60 min of tracing, were analyzed according to FIGO (The International Federation of Gynecology and Obstetrics) guidelines [16] using Omniview-SisPorto, version 4.0.9 [17] and the following FHR features, from the last 60 min of tracings, were registered: FHR baseline, which is the mean level of the most horizontal and less oscillatory FHR segments; percentage of abnormal short-term variability (%abSTV), subsequent FHR signals differing < 1 bpm; percentage of abnormal long term variability (%abLTV), FHR signals with difference between minimum and maximum values in surrounding 1-min window < 5 bpm; mean short-term variability (mean STV); number of acceleration (Acc), i.e., abrupt increases in FHR above the baseline, of more than 15 bpm in amplitude, and lasting more than 15 s, but less than 10 min per minute; number of decelerations (Dec), i.e., decreases in the FHR below the baseline, of more than 15 bpm in amplitude, and lasting more than 15 s per minute.

Newborn umbilical artery blood (UAB) pH was used as measure of fetal oxygenation, as it represents an active measure of fetal oxygenation. A low UAB pH indicates the presence of

acidemia occurring during labor and delivery, presenting a higher risk of perinatal death or neurological injuries from hypoxia.

Of the 68 cases, 48 delivered fetuses with pH in the normal range, pH $\geq$ 7.20 (N), 10 delivered with UAB pH between 7.10 and 7.20, mildly acidemic fetuses (MA), and 10 moderate-to-severe acidemic fetuses with UAB pH $\leq$ 7.10 (MSA).

All of the tracings were resampled at a frequency of 2Hz after pre-processing, based on an algorithm described in previous studies [18]. A more detailed description of the data is presented elsewhere [5].

Spearman Correlation Coefficients were used to compare each complexity measure with different parameters. For entropy, two measures were used: Approximate Entropy (ApEn) and Sample Entropy (SampEn), and for each, 3 different tolerances were used (0.10, 0.15 and 0.2). For compression, six different compressors were used, namely brotli, bzip2, gzip, paq8l, ppmd, and lzma. For the first five, the lowest and highest levels of compression were tested. The latter, lzma, only one level of compression was possible. Compression was measured as compression rate, the compressed size of the trace divided by the original size of the same trace.

Regarding the physiological features that were captured by SisPorto, Spearman Correlation Coefficients between the percentage of abnormal short-term variability (%abSTV), mean value of the STV, baseline, percentage of abnormal long term variability (%abLTV), number of Accelerations (Acc) per 10 min, and Decelerations (Dec) per 10 min were computed.

Spearman Correlation Coefficients between the physiological features and Compressors and between the physiological features and entropies were computed.

Multiscale analysis was also performed to study the effect of entropy and compression, up to scale 20, in the three different groups (N, MA, and MSA). For entropy, tolerance 0.2 was used in both Approximate Entropy and Sample Entropy, while in compression paq8l, brotli, and gzip, with maximum level of compression were used.

Results

The 68 tracings were acquired in singleton pregnancies and were over 32 to 60 min long (mean = 55, standard deviation = 7) with gestational age, for groups MSA (mean = 39.8, standard deviation = 1.3), MA (39.4, 1.6), and N (39.3, 2). Median time between the end of the tracings and the delivery was 0.0 min for all of the groups. The mean (standard deviation) Apgar score at the first minute was 6.2 (2.7) in MSA group, 8.5 (0.7) in MA group, and 8.8 (0.7) in N group.

Correlations between Approximate and Sample entropies with different tolerances were almost perfect, as the lowest Spearman Correlation Coefficient was 0.918. All were significant at the 0.01 level (2-tailed).

As for compression, six different compressors were applied, with five of them having two versions, one as the lowest level of compression (faster) and another with highest level of compression (slower). All of the Spearman Correlation Coefficients are over 0.819. All correlations are significant at the 0.01 level (2-tailed).

Table 3.1 shows the correlations between different compressors (for distinct levels of compression) with entropies (with distinct tolerances). Paq8l_8 has the highest correlation with entropies with Spearman's Coefficient over 0.549 ($p < 0.01$). The lowest correlation coefficients were found between Bzip2 and entropies, as values were under 0.143, with no statistically significant difference found.

Table 3.1. Spearman Correlation Coefficients between different entropies and compressors using different values for tolerance and levels of compression. (ApEn: Approximate entropy; SampEn: Sample Entropy).

	ApEn (0.1)	ApEn (0.15)	SampEn (0.2)	SampEn (0.1)	SampEn (0.15)	SampEn (0.2)
Brotli_1	0.285*	0.267*	0.298*	0.341**	0.294*	0.331**
Brotli_11	0.506**	0.486**	0.508**	0.543**	0.496**	0.531**
Gzip1	0.407**	0.393**	0.424**	0.452**	0.406**	0.448**
Gzip_9	0.242*	0.227	0.258*	0.285*	0.24*	0.281*
Bzip2_1	0.093	0.072	0.105	0.143	0.09	0.132
Bzip2_9	0.093	0.072	0.105	0.143	0.09	0.132
Ppmd_2	0.256*	0.247*	0.28*	0.293*	0.247*	0.288*
Ppmd_16	0.172	0.152	0.182	0.221	0.17	0.21
Paq8l_1	0.531**	0.511**	0.538**	0.57**	0.52**	0.559**
Paq8l_8	0.573**	0.549**	0.573**	0.606**	0.556**	0.592**
Lzma_6	0.355**	0.356**	0.382**	0.366**	0.331**	0.36**

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Table 3.2 describes the correlation between physiological features captured by SisPorto, using

the FIGO guidelines. The highest correlations were found between %abSTV and Mean STV (r = −0.796), and also between %abSTV and %abLTV (r = 0.706). Note that %abSTV was the only variable significantly correlated with all others.

Table 3.2. Spearman Correlation Coefficients between physiological features captured by SisPorto, using the The International Federation of Gynecology and Obstetrics (FIGO) guidelines [1]. (%abSTV: percentage of abnormal; mean STV: mean short-term variability; %abLTV: percentage of abnormal long; Acc: number of acceleration; Dec: number of decelerations).

	%abSTV	Mean STV	%abLTV	Baseline	Acc	Dec
%abSTV	1					
Mean STV	−0.796**	1				
%abLTV	0.706**	−0.444**	1			
baseline	0.324**	−0.201	0.27*	1		
Acc	−0.480**	0.357**	−0.651**	−0.138	1	
Dec	−0.375**	0.539**	−0.101	−0.015	−0.035	1

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Both entropy and compression captured complexity in the tracings, as seen in Tables 3.3 and 3.4, but compression was correlated with more physiologic features. Actually, there were significant correlations between compressors and number of accelerations and decelerations per minute, but neither were correlated with entropies. Despite that there was a significant fair correlation between entropies and %abSTV, its magnitude was rather median, when compared with compression (Tables 3.3 and 3.4).

Table 3.3. Spearman Correlation Coefficients between physiological features and Entropies.

	ApEn (0.1)	ApEn (0.15)	ApEn (0.2)	SampEn (0.1)	SampEn (0.15)	SampEn (0.2)
%abSTV	−0.557**	−0.541**	−0.561**	−0.617**	−0.586**	−0.624**
Mean STV	0.34**	0.331**	0.353**	0.339**	0.302*	0.338**
%abLTV	−0.319**	−0.321**	−0.328*	−0.451**	−0.441**	−0.459**
baseline	−0.268*	−0.29*	−0.284*	−0.312**	−0.309*	−0.313**
Acc	0.093	0.077	0.080	0.199	0.178	0.206
Dec	−0.024	−0.020	0.029	−0.058	−0.103	−0.072

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Table 3.4. Spearman Correlation Coefficients between physiological features and Compressors.

	Brotli_1	Brotli_11	Gzip_1	Gzip_9	Bzip2_1	Bzip2_9	Ppmd_2	Ppmd_16	Paq8l_1	Paq8l_8	Lzma_6
%abSTV	−0.886**	−0.934**	−0.908**	−0.851**	−0.796**	−0.796**	−0.838**	−0.808**	−0.935**	−0.931**	−0.882**
Mean STV	0.783**	0.829**	0.794**	0.774**	0.729**	0.729**	0.774**	0.733**	0.8**	0.783**	0.839**
%abLTV	−0.622**	−0.6**	−0.625**	−0.575**	−0.532**	−0.532**	−0.591**	−0.539**	−0.646	−0.606	−0.556**
baseline	−0.236	−0.295*	−0.249*	−0.193	−0.168	−0.168	−0.288*	−0.175	−0.3*	−0.296*	−0.359*
Acc	0.564**	0.472**	0.542**	0.562**	0.574**	0.574**	0.548**	0.570**	0.527**	0.485**	0.432**
Dec	0.410**	0.387**	0.397**	0.424**	0.441**	0.441**	0.487**	0.423**	0.366**	0.339**	0.544**

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

A multiscale analysis of the tracings using either entropy or compression was also performed. Using Approximate Entropy with tolerance of 0.2, it was possible to distinguish MSA tracings from the other two groups up to scale 7. With Sample Entropy (0.2) the distinction was shown up to scale 4 (Figures 3.1 and 3.2).

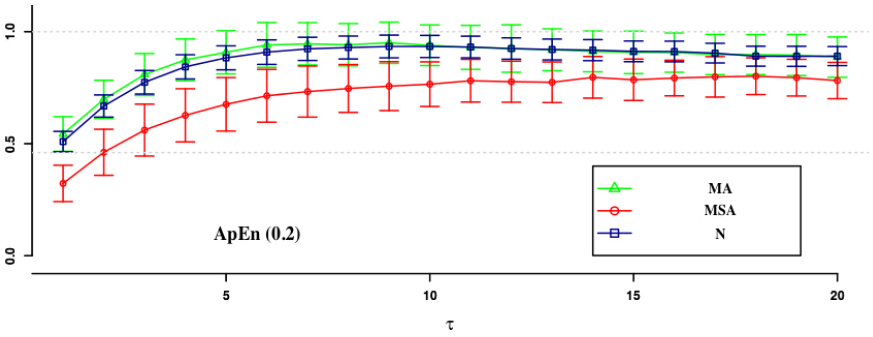

Figure 3.1. Multiscale analysis of tracings using Approximate entropy (ApEn) with tolerance 0.2. Plotted 95% Confidence Intervals of the mean for each group in each scale. (MA: mildly acidemic fetuses; MSA: moderate-to-severe acidemic fetuses; N: the normal range).

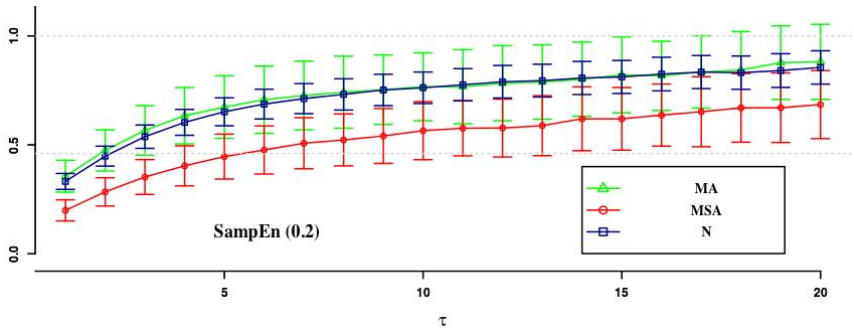

Figure 3.2. Multiscale analysis of tracings using Sample entropy (SampEn) with tolerance 0.2. Plotted 95% Confidence Intervals of the mean for each group in each scale.

Regarding compression, the opposite happens. Using all of the compressors, it was possible to distinguish groups N and MA in lower scales. Actually, paq8l distinguished them in all scales used. In all compressors, as scale increases, group N starts to diverge from group MSA, with bzip2 and gzip getting statistically significant results for scale higher than 2 and 7, respectively (Figures 3.3–3.6).

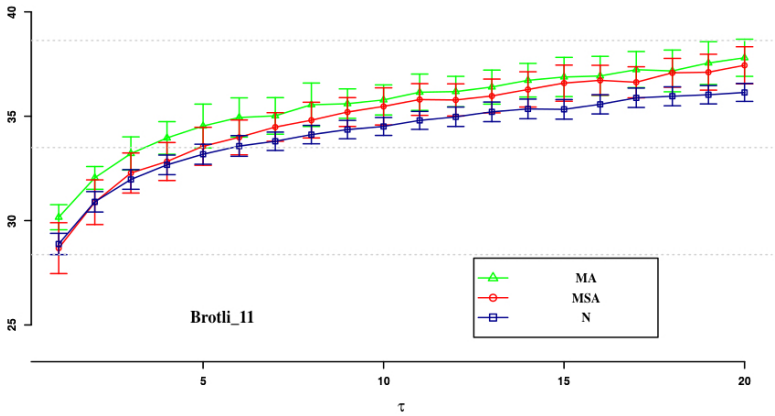

Figure 3.3. Multiscale analysis of tracings using compressor Brotli with maximum level of compression. Plotted 95% Confidence Intervals of the mean for each group in each scale.

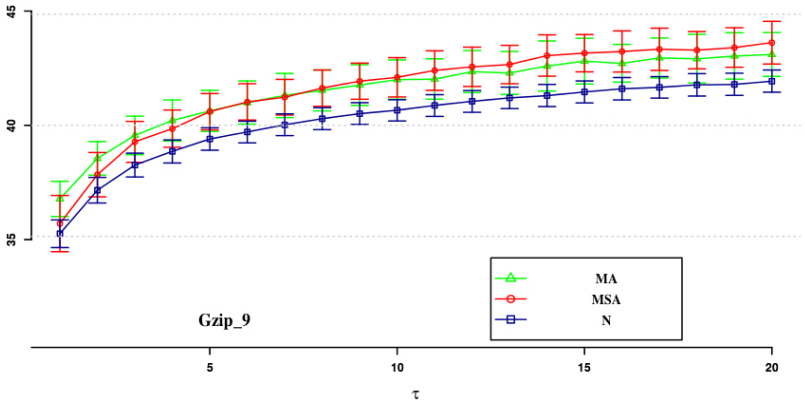

Figure 3.4. Multiscale analysis of tracings using compressor Gzip with maximum level of compression. Plotted 95% Confidence Intervals of the mean for each group in each scale.

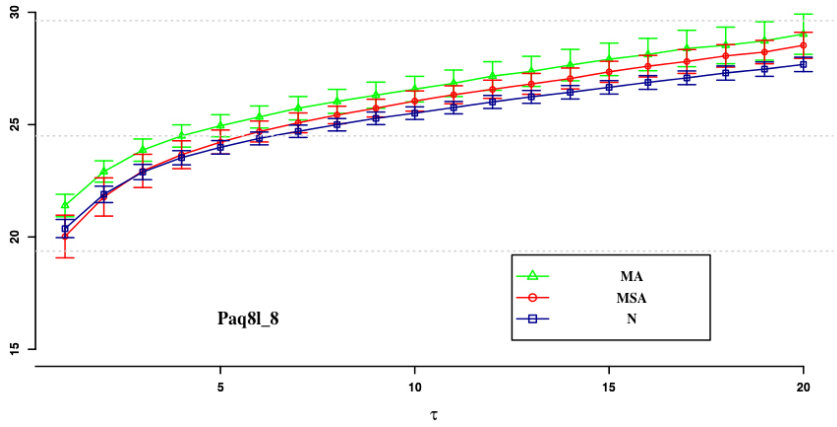

Figure 3.5. Multiscale analysis of tracings using compressor Paq8l with maximum level of compression. Plotted 95% Confidence Intervals of the mean for each group in each scale.

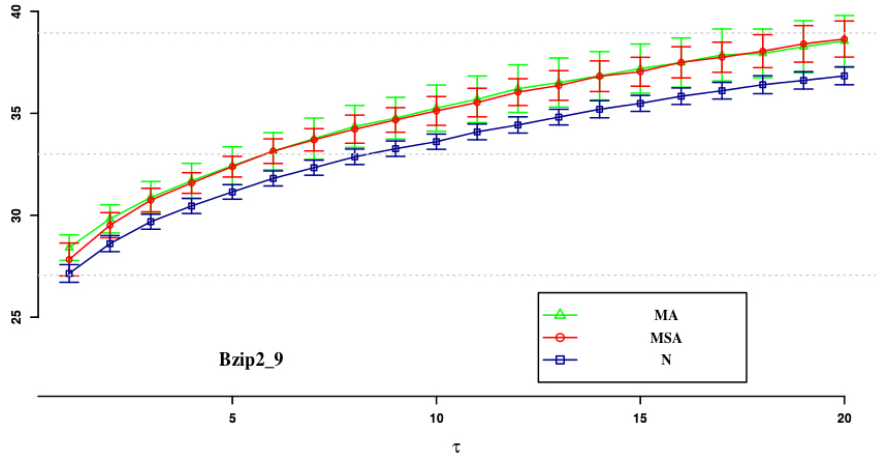

Figure 3.6. Multiscale analysis of tracings using compressor Bzip2 with maximum level of compression. Plotted 95% Confidence Intervals of the mean for each group in each scale.

Discussion

In this study, the different entropies and tolerances used in FHR analysis were highly correlated as were the different used compressors. This suggests that different entropy and compressor estimators are internally valid, i.e., despite some discrepancies in results, they seem to absorb similar information from the tracings.

Despite the high correlation within similar complexity measures, the use of different values for certain parameters, such as the threshold r when using ApEn, might lead to different performances in the characterization of different fetal behavioral patterns or acute and chronic conditions [20]. Besides, different complexity measures seem to capture different information since correlation between entropies and compressors is low, with bzip2 being perfectly uncorrelated, while ppmd and gzip having some non-significant correlations with entropies. On the other hand, paq8l is moderately correlated. It is still not totally clear why this happens, but maybe some characteristics of each compressor algorithm may provide us some answers. For example, in order to optimize the compression process, bzip2 performs a block sort (reversible), and, consequently, does not take advantage from the initial structure (related with entropy).

As to the main objective of this study of comparing the two different complexity measures, entropy and compression, with the FIGO international guidelines for fetal monitoring, we have observed a good correlation between percentage of abnormal short-term variability and entropy. However, compression has seemed to be the measure that has captured most of the information out of abnormal short-term variability. Higher values of entropy and compression mean that the presence of more complex structures and these structures were characteristic from healthier status. Tracings with lower percentage of abnormal STV were typical from healthy babies, which explains the negative correlation between the variables (compression/entropy and abnormal short-term variability).

There is no uniform definition of STV. The one that is used in this work was defined in SisPorto as subsequent FHR signals differing < 1 bpm, closely following the FIGO guidelines, but other definitions exist in the literature, with no concrete agreement on which one should be used [21]. It would be interesting to see how compression behaves with these other variants of

STV. In an affirmative case, compression could be used as a universal measure to capture STV. Moreover, compression seems to capture more features from tracings since numbers of accelerations and decelerations per minute were correlated with compression but not with entropies.

Actually, it is interesting to notice that the accelerations and decelerations did not correlate with any entropy but correlated with all complexities. As in Baumert's study [22], our results suggest that acceleration and decelerations are not purely random but follow some deterministic structures that can be explored by compression algorithms. The use of different parameters in the entropies, considering for instance variants of the threshold such as fuzzy functions, might allow for capturing accelerations and decelerations.

With multiscale analysis, it was observed that entropy and compression captured distinct clinical information from the tracings. Entropy distinguishes MSA tracings from the other two groups based on some features that compression cannot capture, and, on the other hand, compression distinguishes groups N and MA possibly by using all SisPorto features. Probably, entropy is correlated with some other features (such as the dynamics) of the physiological data that is not captured by SisPorto. As in Voss study [23] these results show that several nonlinear indices should be combined in order to improve the performance of FHR analysis. Moreover, lower scales in entropy and higher scales in compression can distinguish groups (MSA, MA and N). Compression is better correlated with physiological features that are captured by SisPorto system, and, in fact, these physiological features also do not distinguish groups at scale one. On the other hand, entropy is able to distinguish groups in lower scales, probably because entropy is correlated with some other features, such as the dynamics of the physiological data that are not captured by SisPorto. Notice that, in higher scales, the averaging of the points can justify the absence of these dynamics, lowering the ability to distinguish the classes.

One limitation of this work is the low number of cases, particularly in acidemic groups. This is particularly due to the low prevalence of these cases, which limits the obtention of datasets with higher sample size. It would be interesting, as a future work, and with a bigger dataset, to try to build a predictive model and evaluate how much better would entropy/compression plus SisPorto features perform compared to SisPorto features alone.

Our results enhance the idea that entropy and compression are complementary complexity measures. More research in this area should be done, regarding higher scale values (with long duration tracings), and the possibility of building a model combining both measures.

Acknowledgments:

João Monteiro-Santos acknowledges the support of the Digi-NewB project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689260.
Luís Antunes, Mohammad Nozari and Cristina Costa-Santos acknowledge the support of Project "NanoSTIMA: Macro-to-Nano Human Sensing: Towards Integrated Multimodal Health Monitoring and Analytics/NORTE-01-0145-FEDER-000016" is financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF).
Hernâni Gonçalves acknowledges the support by ERDF through the operation POCI-01-0145-FEDER-007746 funded by the Programa Operacional Competitividade e Internacionalização – COMPETE2020 and by National Funds through FCT - Fundação para a Ciência e a Tecnologia within CINTESIS, R&D Unit (reference UID/IC/4255/2013).

Author Contributions:

João Monteiro-Santos and Cristina Costa-Santos substantial contribution to conception and design; João Monteiro-Santos, Cristina Costa-Santos, João Bernardes and Hernâni Gonçalves for interpretation of data; Mohammad Nozari analyzed data; João Bernardes, Hernâni Gonçalves, Luís Antunes, Cristina Costa-Santos revise critically for important intellectual content; João Monteiro-Santos and Cristina Costa-Santos wrote the paper.

Conflicts of Interest:

João Bernardes has been involved in the development of the commercially available SisPorto system for FHR monitoring.

References

[1]. Ayres-de-Campos, D.; Bernardes, J.; Costa-Pereira, A.; Pereira-Leite, L. Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. Br. J. Obstet. Gynaecol. 1999, 106, 1307–1310.

[2]. Bernardes, J.; Ayres-de-Campos, D. The persistent challenge of fetal heart rate monitoring. *Curr. Opin. Obstet. Gynecol.* **2010**, *22*, 104–109.

[3]. Bernardes, J.; Costa-Pereira, A.; Ayres-de-Campos, D.; van Geijn, H.P.; Pereira-Leite, L. Evaluation of interobserver agreement of cardiotocograms. *Int. J. Gynaecol. Obstet.* **1997**, *57*, 33–37.

[4]. Nunes, I.; Ayres-de-Campos, D.; Figueiredo, C.; Bernardes, J. An overview of central fetal monitoring systems in labor. *J. Perinat. Med.* **2013**, *41*, 93–99.

[5]. Goncalves, H.; Rocha, A.P.; 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*, 847–855.

[6]. Ferrario, M.; Signorini, M.G.; Magenes, G. Complexity analysis of the fetal heart rate variability: Early identification of severe intrauterine growth-restricted fetuses. *Med. Biol. Eng. Comput.* **2009**, *47*, 911–919.

[7]. Plsek, P.E.; Greenhalgh, T. Complexity science: The challenge of complexity in health care. *BMJ* **2001**, *323*, 625–628.

[8]. Wilson, T.; Holt, T.; Greenhalgh, T. Complexity science: Complexity and clinical care. *BMJ* **2001**, *323*, 685–688.

[9]. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301.

[10]. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049.

[11]. Cilibrasi, R.; Vitanyi, P.M.B. Clustering by compression. *IEEE Trans. Inf. Theory* **2005**, *51*, 1523–1545.

[12]. Cilibrasi, R.; Vitányi, P.; Wolf, R.D. Algorithmic clustering of music based on string compression. *Comput. Music J.* **2004**, *28*, 49–67.

[13]. Wehner, S. Analyzing worms and network traffic using compression. *J. Comput. Secur.* **2007**, *15*, 303–320.

[14]. Santos, C.C.; Bernardes, J.; Vitanyi, P.M.B.; Antunes, L. Clustering fetal heart rate tracings by compression. In Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems, Salt Lake City, UT, USA, 22–23 June 2006; IEEE Computer Society: Washington, DC, USA, 2006; pp. 685–690.

[15]. Henriques, T.; Goncalves, H.; Antunes, L.; Matias, M.; Bernardes, J.; Costa-Santos, C. Entropy and compression: Two measures of complexity. *J. Eval. Clin. Pract.* **2013**, *19*, 1101–1106.

[16]. Ayres-de-Campos, D.; Spong, C.Y.; Chandraharan, E.; Panel, F.I.F.M.E.C. Figo consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *Int. J. Gynaecol. Obstet.* **2015**, *131*, 13–24.

[17]. Ayres-de-Campos, D.; Sousa, P.; Costa, A.; Bernardes, J. Omniview-sisporto® 3.5—A central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis. *J. Perinat. Med.* **2008**, *36*, 260–264.

[18]. Goncalves, H.; Rocha, A.P.; Ayres-de-Campos, D.; Bernardes, J. Internal versus external intrapartum fetal heart rate monitoring: The effect on linear and nonlinear parameters. *Physiol. Meas.* **2006**, *27*, 307–319.

[19]. 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*, 62–67.

[20]. Goncalves, H.; Rocha, A.P.; 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. Hematol. Disord. Drug Targets* **2008**, *8*, 91–98.

[21]. Cesarelli, M.; Romano, M.; Bifulco, P. Comparison of short-term variability indexes in cardiotocographic fetal monitoring. *Comput. Biol. Med.* **2009**, *39*, 106–118.

[22]. Baumert, M.; Baier, V.; Voss, A.; Brechtel, L.; Haueisen, J. Estimating the complexity of heart rate fluctuations—An approach based on compression entropy. *Fluct. Noise Lett.* **2005**, *5*, L557–L563.

[23]. Voss, A.; Schulz, S.; Schroeder, R.; Baumert, M.; Caminal, P. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. A Math. Phys. Eng. Sci.* **2009**, *367*, 277–296.

COMPLEXITY OF CARDIOTOCOGRAPHIC SIGNALS AS A PREDICTOR OF LABOR

JOÃO SANTOS, CRISTINA SANTOS, TERESA HENRIQUES, JOÃO BERNARDES, INÊS NUNES, CÉLIA COSTA

ENTROPY

SPECIAL ISSUE "ASSESSING COMPLEXITY IN PHYSIOLOGICAL SYSTEMS THROUGH BIOMEDICAL SIGNAL ANALYSIS

JANUARY 2020

4. Complexity of Cardiotocographic Signals as A Predictor of Labor

Abstract

Prediction of labor is of extreme importance in obstetric care to allow for preventive measures, assuring that both baby and mother have the best possible care. In this work, the authors studied how important nonlinear parameters (entropy and compression) can be as labor predictors. Linear features retrieved from the SisPorto system for cardiotocogram analysis and nonlinear measures were used to predict labor in a dataset of 1072 antepartum tracings, at between 30 and 35 weeks of gestation. Two groups were defined: Group A—fetuses whose traces date was less than one or two weeks before labor, and Group B—fetuses whose traces date was at least one or two weeks before labor. Results suggest that, compared with linear features such as decelerations and variability indices, compression improves labor prediction both within one (C-Statistics of 0.728) and two weeks (C-Statistics of 0.704). Moreover, the correlation between compression and long-term variability was significantly different in groups A and B, denoting that compression and heart rate variability look at different information associated with whether the fetus is closer to or further from labor onset. Nonlinear measures, compression in particular, may be useful in improving labor prediction as a complement to other fetal heart rate features.

Introduction

Worldwide, approximately 15 million infants are born preterm (after less than 37 completed weeks of gestation) each year [1]. Over one-third of the world's estimated 3 million annual neonatal deaths are related to preterm birth [2–4]. Even after surviving the neonatal period, infants born preterm are at increased risk of delayed childhood development and low economic productivity [5]. Therefore, interventions to reduce the preterm birth rate are of utmost importance.

Clinical decisions during labor and delivery in developed countries are strongly based on cardiotocography (CTG) [6–8], which has been one of the most used tools in assessing fetal well-being since the early ’60s. CTG combines fetal heart rate (FHR), obtained using a Doppler ultrasound probe or electrocardiogram electrodes, with uterine contractions (UC) measurements, obtained using an abdominal or intra-uterine pressure transducer. Both provide relevant information about the fetal condition and early detection of preterm labor and abnormal labor progress [7,9,10].

Despite the importance of assessing the well-being of the fetus and mother, poor agreement among physicians in the analysis and classification of CTGs is still a problem, even among experienced obstetricians, resulting in a high false positive rate [6,11,12]. In daily practice, FHR and UC are displayed on a printout or monitor to be visually interpreted by a clinician. Even when following specific, well-accepted guidelines (for example, the International Federation of Obstetrics and Gynecology (FIGO), associated with high sensitivity and low specificity [13]), interpretation of CTG relies on the clinician’s opinion and daily practice. This leads to a chance that adherence to conventional guidelines could be more harmful than beneficial [14].

The beat-to-beat variation of FHR reflects the influence of the fetus’ autonomic nervous system (ANS) and its components (sympathetic and parasympathetic) in the heart. Therefore, it is an indicator of the fetal pathophysiological status, which can be used in the assessment of fetal well-being [15] and its well-known influence on labor onset and progression [16]. 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 [17]. The linear modeling approach is used to quantify sympathetic and parasympathetic control mechanisms and their balance through the measurement of 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 [18]. For this matter, in the past couple of decades, and with the fast development of computation, new signal processing and pattern recognition methodologies (namely entropy and compression) have been developed and applied to many different fields, including the analysis of fHRV

[19,20]. These approaches can reveal relevant clinical information not exposed by temporal or frequency analysis [21].

Systems, such as Omniview SisPorto [22–24] and NST-Expert, which later became CAFE [25], can automatically deal with CTG assessment and then overcome the limitations of the visual assessment of CTGs mentioned above, but clinical judgment remains highly dependent on CTG analysis [26]. Since all FHR processing and analysis in these systems is based on morphological features provided by FIGO guidelines, they lack the integration of nonlinear indices that would allow them to be optimized.

The ability to predict preterm labor can improve the well-being of both fetus and mother. The successful prediction of preterm labor is an essential part of a decision support system for physicians to implement measures that adequately reduce related fetal morbidity and mortality (like the administration of corticosteroids to the mother in order to accelerate lung maturation and therefore decrease the risk of respiratory distress in the newborn).

The main objective of this work is to evaluate how useful nonlinear parameters, namely entropy and compression, can be as labor predictors by using antepartum FHR and UC traces one or two weeks before labor.

Methods

Nonlinear Methods

Compression

The Kolmogorov Complexity (KC) [27] is defined as the function mapping a string x in an integer, bounded to a Turing Machine ϕ. The KC reflects the increase in new patterns along a given sequence. In this case, the word complexity refers to the algorithmic complexity, defined according to information theory, as the length of the shortest program p able to print the string x.

$$KC_{\phi}(x) = \begin{cases} min\{|p|: \phi(p) = x\}, & if\ \phi(p) = x \\ \infty & if\ p\ does\ not\ exit, \end{cases} \quad (1)$$

For a random string, the output of the KC function will be the length of the original string, as any compression effort will end in information loss. On the other hand, the more reoccurring patterns, the less complex the string is.

Although this concept is objective, its applicability is limited by the fact that KC is not computable. Compressors are a close upper-bounded approximation of the KC function. For over 30 years, data compression software has been developed for data storage and transmission efficiency purposes. More recently, compression has been utilized in research fields like music, literature, internet traffic, and health [28–30].

In this work, we will assess the algorithmic complexity of FHR and UC signals by applying the Gzip compressor. Gzip [31] combines two classical algorithms—Lempel–Ziv (LZ77) [32], a dictionary-based algorithm, and Huffman scheme [33]—by encoding sequences of high probability using shorter bits in comparison with lower probability strings, where longer bits are used. The amount of compression obtained depends on the input file size and the distribution of common substrings.

The idea is that for a given time series, the compression ratio (CR), i.e., the compressed size of the file divided by its original size, can be used to assess the complexity. A random series will have CR close to 1, whereas a series full of patterns will be highly compressible and, therefore, the CR will be close to 0. The Gzip with maximum compression levels and values presented represents the percentage of CR.

Entropy

In 1991, Pincus developed the Approximate Entropy (ApEn), a regularity statistic tool used to quantify a system's complexity based on the notion of entropy [34]. The ApEn measures the

irregularity of time series and is defined as the logarithmic likelihood that the patterns of a time series that are close to each other will remain close when longer patterns are compared.

Later, in 2000, Richman and Moorman [35] proposed Sample Entropy (SampEn). Similar to ApEn, the SampEn measures time series irregularity. However, it does so with some major advantages: 1) self-matches are not counted, reducing bias; 2) it agrees much better than ApEn statistics with the theory for random numbers with known probabilistic character over a broad range of operating conditions; 3) the conditional probabilities are not estimated in a template manner. Instead, they are computed directly as the logarithm of conditional probability rather than from the ratio of the logarithmic sums, showing relative consistency in cases where ApEn does not [36].

To use either ApEn or SampEn, decisions on two different parameters, m, and r, have to be made. The m parameter is the embedding dimension, i.e., the length of sequences to be compared, while the tolerance parameter r works as a similarity threshold. Two patterns are considered similar if the difference between any pair of corresponding measurements is less than or equal to r. Values of 0.1, 0.15, or 0.2 standard deviations (SD) are usually used for parameter r, while m is mostly considered as 2 [37]. In this work, tolerance of 0.1 SD and an embedding dimension of 2 were used.

Data

The FHR data used for this study were from a retrospective cross-sectional study [38]. Each FHR trace corresponds to distinct fetuses from a singleton pregnancy. The selected traces were acquired between July 2005 and November 2010 during hospitalization in a tertiary care university hospital. All traces were acquired at least 48 h before delivery to guarantee they included no labor time. Furthermore, the traces included were at least 20 min long, during which the signal quality was over 80%, and the signal loss was less than 33%.

The cardiotocographic signals were acquired using an external ultrasound sensor applied to the maternal abdomen. The ultrasound signal is filtered, envelope rectified and digitized at a sampling rate of 800 Hz with a 12-bit precision [39]. Then, an autocorrelation function is used

to calculate the heart period and the similarity between pulses of two consecutive heartbeats, as described in [40]. Via the digital outputs of the fetal monitors, resulting traces were analyzed using the Omniview SisPorto® 3.7 system [23] at a sampling rate of 4 Hz (Figure 4.1).

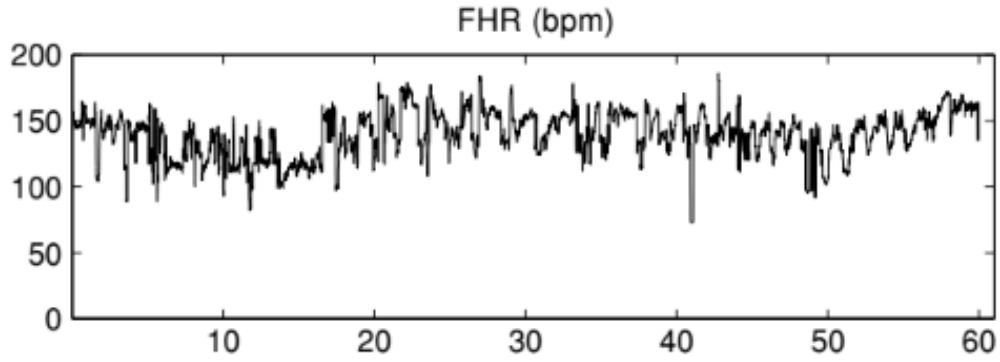

Figure 4.1. Example of a fetal heart rate (FHR) time series.

SisPorto features used in this paper are summarily described in Table 4.1. Note that the SisPorto system does not perform any average or reduction in FHR/UC signals.

The 1072 traces selected ranged from 30 to 35 gestational weeks. Two groups were defined: Group A—fetuses whose traces date was less than two weeks before labor, and Group B—fetuses whose traces date was at least two weeks before labor. Physiological fetal and maternal features, such as maternal age (mAge) and baby gender, as well as some tracing characteristics such as trace duration and signal quality, were compared in both groups. Linear indices for uterine contraction analysis comprised of mean_UC (median of UC mean from 10min nonoverlapping blocks), sd_UC (median of UC standard deviation from 10min nonoverlapping blocks) and cv_UC (coefficient of variability of UC).

Two complexity measures, Gzip and SampEn, were considered in this work. Because the value of these measures depends on the trace size, each tracing was split into non-overlapping blocks of 10 min. Both Gzip and SampEn were computed for each block. Then, the median value of CR and SampEn for each fetus was used. Both complexity measures were calculated for FHR (Gzip_FHR and SampEn_FHR) and UC signals (Gzip_UC and SampEn_UC).

Table 4.1. Description of SisPorto features [22,24].

SisPorto Variable	Description
Basal line FHR	mean level of the most horizontal and less oscillatory FHR segments, in the absence of fetal movements and uterine contraction (UC), associated with periods of fetal rest, estimated via a complex algorithm
baseline	approximation of basal FHR to long-term FHR fluctuations using running averaging
number of accelerations (nAccel)	number of increases in FHR over the baseline lasting 15–120 s and reaching a peak of at least 15 bpm in 60 min
number of contractions (nContr)	number of periods in 60 min, lasting a maximum of 254 s, where an upward slope exceeding 17 s was detected reaching a peak lasting more than 90 s, followed by a downward slope exceeding 17 s
number of mild decelerations (mDec)	number of decreases in FHR under the baseline lasting 15–120 s, with a minimum amplitude of 15 bpm in 60 min
number of intermediate decelerations (iDec)	number of decreases in FHR under the baseline lasting 120–300 s, with a minimum amplitude of 15 bpm in 60 min
number of prolonged decelerations (pDec)	number of decelerations lasting more than 300 s in 60 min
average short-term variability (avSTV)	mean difference between adjacent FHR signals at 4 Hz on the fetal monitor, after removal of adjacent signals that differ >15 bpm
abnormal short-term variability (abSTV)	percentage of subsequent FHR signals differing <1 bpm
average long-term variability (avLTV)	mean difference between max and min FHR in a 1 min sliding window, in segments free of accelerations or deceleration
abnormal long-term variability (abLTV)	percentage of FHR signals with a difference between minimum and maximum values in a surrounding 1 min window <5 bpm

Statistical Analysis

Normality for continuous variables was evaluated by visual inspection of the frequency distribution (histogram). For normally distributed variables, the values for each group are presented as mean ± SD, and an independent samples t-test was performed. On the other hand, for skewed continuous variables, the values are presented as median (minimum-maximum), and the Mann–Whitney test was used to compare the two groups. The categorical variables were compared in the two groups applying the Chi-Square test or Fisher's exact test as applicable.

Logistic regression, using Hosmer–Lemeshow to test the goodness of fit, was used to predict which fetuses will be born preterm in the next two weeks. Variables were selected using Wald's backwards method. The concordance statistic (C-statistic), measured by the area under the receiver operating characteristic curve, was computed to assess the model's discrimination.

Akaike Information Criterion (AIC), $AIC = 2k - 2\log(L)$, where k is the number of parameters and L the maximum value of the likelihood function, was used for model comparison, where a lower result suggests a better model.

Statistical analysis was performed with IBM SPSS Statistics for Windows, version 24 (IBM, Armonk, NY, USA).

Results

A total of 1072 antepartum tracings were used, 96 of which were born in the following two weeks (Group A). The main clinical characteristics of the group in which fetuses were born in the next two weeks (Group A) and the group in which they were not (Group B) are presented and compared in Table 4.2. Note that no differences were found between the groups for these variables.

Table 4.2. Fetal and maternal features from Group A–fetuses whose traces date was less than two weeks before labor, and Group B–fetuses whose traces date was at least two weeks before labor.

	Group A (n = 96) Median (min-max), Mean ± SD or N (%)	Group B (n = 976) Median (min-max), Mean ± SD or N (%)	P-Value
Trace duration (min)	25.56 (14.82–67.07)	25.18 (11.28–96.31)	0.905
Gestational age at delivery (weeks)	36.58 ± 1.12	38.92 ± 1.20	
Maternal age (years)	31 (16–43)	31 (15–52)	0.291
Cesarean section	31 (32.3)	321 (32.9)	0.067
Baby presentation (cephalic)	90 (93.8)	918 (94.1)	0.524
Gender (male)	49 (51)	506 (51.8)	0.881
Signal quality (%)	97 (80–100)	96 (80–100)	0.105
Signal loss (%)	3 (0–20)	4 (0–21)	0.106

SisPorto features were also compared between the two groups (Table 4.3). Statistical significance was found with variables iDec ($p < 0.001$), which was lower in fetuses who would be born in the next two weeks, and average long-term variability (abLTV), which was higher in fetuses who would be born in the next two weeks ($p = 0.038$).

Furthermore, while SampEn was not able to find differences between the traces from babies in the two groups with FHR and UC signals, Gzip was ($p = 0.024$ for FHR, $p = 0.013$ for UC), being lower in fetuses who would be born in the next two weeks (Group A) for FHR signals, while the opposite happened for UC signals. The standard deviation of UC was also significantly higher for Group A ($p = 0.020$).

Table 4.3. SisPorto and nonlinear features from Group A–fetuses whose traces date were less than two weeks before labor, and Group B –fetuses whose traces date were at least two weeks before labor.

	Group A (n = 96) Median (min-max), Mean ± SD or N (%)	Group B (n = 976) Median (min-max), Mean ± SD or N (%)	P-Value
Basal line	133 (108–154)	134 (105–168)	0.137
Baseline	135.5 (114–160)	137 (105–169)	0.237
nAccel	5 (0–13)	5 (0–31)	0.188
nContr	1 (0–15)	1 (0–15)	0.200
mDec	0 (0–5)	0 (0–13)	0.787
iDec (% of no iDec)	89 (92.71)	962 (98.57)	**<0.001**
pDec (% of no pDec)	96 (100)	973 (99.69)	1.000
abSTV	50.49 ± 8.83	50.27 ± 8.42	0.805
avSTV	14.48 ± 3.48	14.55 ± 3.45	0.839
abLTV	1 (0–35)	0 (0–38)	**0.038**
avLTV	15.85 (8–33)	16.8 (0–40)	0.229
mean_UC	172.504 ± 103.426	166.663 ± 101.650	0.592
sd_UC	56.350 ± 42.403	45.768 ± 35.096	**0.020**
cv_UC	0.424 ± 0.347	0.369 ± 0.328	0.121
Gzip_UC	6.089 ± 1.769	5.664 ± 1.568	**0.013**
SampEn_UC	0.547 ± 0.306	0.595 ± 0.287	0.117
Gzip_FHR	11.559 ± 0.995	11.758 ± 0.878	**0.024**
SampEn_FHR	0.670 ± 0.159	0.693 ± 0.195	0.265

Logistic regression, including all relevant variables ($p < 0.05$)—Gzip_FHR, Gzip_UC, sd_UC, iDec, a week of CTG (wCTG), and abLTV—was then performed using a backward selection model. The model obtained included the variables Gzip, iDec and a week of CTG (wCTG). Also, interactions between Gzip and wCTG were considered but found to be non-significant. Results from the logistic regression can be found in Table 4.4.

Table 4.4. Logistic regression for labor prediction in two weeks or less.

	B	P-value	Exp(B)	95% CI
Constant	−20.639	<0.001		
wCTG	0.674	<0.001	1.962	1.489–2.584
Gzip_FHR	−0.341	0.005	0.711	0.560–0.902
iDec[a]	1.782	<0.001	5.950	2.217–15.918

[a] No iDec was set as reference instance.

From this logistic regression model, abLTV and UC variables were removed from the initial set of predictors made by the model, and a C-statistic of 0.704 was obtained, with a 95% confidence interval range of 0.651–0.758. Also, the AIC obtained for this model was 603.763. The process was repeated considering all relevant physiological and linear features but

without Gzip. This model, now without Gzip but with abLTV, achieved an AIC of 605.5 and a C-statistic of 0.691 (0.639–0.742).

The groups were also redefined and tested again. The same analysis as before was performed, except Group A consisted of fetuses who were born less than one week (instead of two weeks) from trace acquisition (n = 27, all preterm) and Group B consisted of all other fetuses (n = 1045, term and preterm babies), which were born as term and preterm babies. SisPorto and nonlinear features were compared between the groups, as carried out in our previous analysis (results in Appendix).

The logistic regression results are shown in Table 4.5. Note that the same variables were included in the logistic regression.

Table 4.5. Logistic regression for labor prediction in one week or less.

	B	P-value	Exp(B)	95% CI
Constant	−6.679	0.330		
wCTG	0.317	0.097	1.373	0.944–1.997
Gzip_FHR	−0.573	0.010	0.564	0.364–0.873
iDec	2.780	<0.001	16.112	5.205–49.874

This model achieved an AIC of 235.3 and a C-statistic of 0.728 (0.619–0.836), which is a small improvement compared with the first one described in this paper.

In Table 4.6, Spearman's correlation coefficient between Gzip and different physiological measures of variability was calculated. Moreover, the same coefficient was calculated for each group. Statistically significant results were found for abLTV and avLTV for two weeks labor prediction.

Table 4.6. Spearman's correlation coefficient and respective 95% confidence interval (CI) between Gzip_FHR and short- and long-term variabilities given by SisPorto. Confidence intervals were calculated using bootstrapping. Bold means significant differences between groups.

		Two Weeks Prediction		One Week Prediction	
	Total	Group A	Group B	Group A	Group B
abSTV	−0.524 (-0.564; -0.481)	−0.636 (−0.733; −0.501)	−0.512 (−0.565; −0.463)	−0.694 (−0.867; −0.370)	−0.515 (−0.560; −0.468)
avSTV	0.500 (0.452; 0.541)	0.596 (0.442; 0.720)	0.489 (0.437; 0.539)	0.698 (0.410; 0.864)	0.492 (0.444; 0.539)
abLTV	−0.562 (−0.602; −0.520)	**−0.722 (−0.807; −0.601)**	**−0.541 (−0.589; −0.495)**	−0.760 (−0.893; −0.489)	−0.551 (−0.596; −0.509)
avLTV	0.765 (0.737; 0.792)	**0.885 (0.818; 0.924)**	**0.751 (0.718; 0.780)**	0.874 (0.663; 0.970)	0.760 (0.730; 0.789)

Discussion

This study enhances the importance of the inclusion of nonlinear indices in clinical practice. In particular, the results suggest that the Gzip compression ratio, a measure of the time series complexity, may improve the predictability of labor onset when applied to FHR and UC signals.

The main objective of this work was to predict labor within two weeks. Both groups included preterm and term babies. In Group A, 46 of 90 were term babies, born between 36 and 37 weeks of gestational age; while in Group B, 44 of 976 fetuses were preterm. No statistical significance was found between term and preterm cases in Group A or Group B.

The information captured by compression relates to the information comprised of other physiological features, such as short and long-term variabilities [41]. In our study, Gzip_FHR has a Spearman's correlation coefficient of −0.524 and 0.5 with abSTV and avSTV's variabilities, respectively. These results contrast with a previous study [41] where correlation values were much higher in absolute value (−0.851 and 0.774). Some different characteristics of the datasets used in each study can explain these differences. On the one hand, the dataset of our study was acquired in an antepartum setting, while the data from the previous study were recorded during the intrapartum. In line with this, the difference observed in the two studies suggests that compression looks at physiological regulatory mechanisms that differ between both settings. On the other hand, another possible explanation is the different sampling rates used in the two studies (4 Hz here, versus 2 Hz in the other study). This may indicate that some information is lost when using 2 Hz. This inkling is supported by the results of Gonçalves et al. [42], who found nonlinear differences between both sampling rates. However, the study of Gonçalves et al. [42] is an intrapartum study, and the tolerance parameter for entropy was computed using an automatic threshold proposed by Lu [43]. A multiscale analysis of scale two would be affected by the latter hypothesis (as it mimics a 2 Hz sampling rate), but in our study, no difference was found. Govindan et al. [44] suggested a different approach, modifying the definition of sample entropy using a time delay. Future studies should compare several methods to study the oversampling question.

When factoring by group, we found significant differences in correlations between Gzip and abLTV and avLTV (Table 4.6). Different studies [45–48] found HRV changes, such as variability

increase and pattern formation throughout fetal maturation, captured by nonlinear indices. Here, different patterns arise in the two groups presented, meaning that compression attains different information from HRV when compared with usual metrics. However, no statistical significance was found in one-week labor prediction analysis. We believe this might be due to low statistical power, as the number of individuals in Group A was 27, making confidence intervals too wide.

Some papers [49,50] indicate different gender development throughout gestation and suggest taking this into account in model creation. Though it was taken into consideration, no significant results were found.

The mean compression ratio (instead of median) of the tracings' block was also considered, and the results obtained were similar. These results suggest robustness of compression regarding skewness and outliers, as well as low intra-tracing variability. Furthermore, multiscale analysis [51] was also performed both for SampEn and Gzip up to five scales, since we were using intervals of 10 min (~1440 data points), but no improvement was found.

Two different definitions for the groups were tested. The same analysis as before was performed, considering Group A as babies who were born preterm less than two weeks, and then less than one week, from trace acquisition (n = 27). As shown in Tables 4 and 5, the logistic regression included the same variables. A small improvement was verified when considering one week, compared with two weeks, from labor. These results reinforce the stability of compression when predicting labor time.

Nonlinear FHR features recognition is a problem in the clinical community because clinicians do not always know how to interpret it. Although entropy has been associated with the activity of central nervous system regulation [52,53], there are still no direct associations between compression and the fetus' physiology. Compression looks for patterns in the series, and a healthy fetus is linked with a high compression ratio (a more chaotic signal leads to fewer patterns that are able to be compressed). In contrast, an unhealthy fetus, under the response of its regulatory system, creates a heart rate signal with more patterns, leading to a lower compression ratio. There is evidence that sympatho-vagal activity, and probably also central nervous system activity, are associated with the onset and progression of labor,

namely via sympathetic activation and vagal inhibition mechanisms [16]. A continuous decrease in the sympathetic stress response during the last weeks before labor was also reported [54], contrary to a stable baseline sympathetic level. Being able to find links between these events and nonlinear indices is key for medical acceptance of these tools in daily practice. Therefore, it is imperative that a more thorough analysis of the FHR changes captured by compression is carried out in particular.

These results are relevant since an early prediction of labor as a decision support system for physicians can improve both fetus and mother assessment and care. In particular, being capable of predicting preterm labor is of extreme importance, as major risks to fetus and mother are associated with it.

This work has some limitations. The number of preterm cases is small, considering the week of the CTG variable is included. Because of this, only fetuses between weeks 30 and 35 of gestational age were selected, limiting the interpretability of the results. Although all the cases were hospitalized, no knowledge of the hospitalization cause is known.

Future studies should validate these models in larger datasets and, if possible, test them in different settings, such as during hospitalization and regular appointments.

Conclusions

Prediction of labor is of extreme importance since physicians will be able to take preventive measures to ensure that both baby and mother will be as prepared as possible. In this work, it was shown that nonlinear measures, compression in particular, can improve labor prediction.

Author Contributions:

J.M.-S., T.H., and C.C.-S. substantial contribution to conception and design; T.H., C.C.-S., J.B., I.N., and C.A.-C. revise critically for important intellectual content; J.M.-S. and T.H. wrote the paper.

Funding:

This work was supported by the project "Digi-NewB" funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689260.

Acknowledgments:

J. Santos acknowledges the support of the doctorate scholarship from the Clinical Research and Health Services Program, funded by the European Social Fund under the Portuguese Norte 2020 research and innovation program (grant agreement n° Norte-08-5369-FSE-000063). The authors also acknowledge the SisPorto project based at the department of Obstetrics and Gynecology of the School of Medicine, University of Porto.

Conflicts of Interest:

João Bernardes currently receives royalties from the development of the commercially available SisPorto system for CTG monitoring.

Appendix

Table 4.A1. SisPorto and nonlinear features. from Group A—fetuses whose traces date were less than one week before labor, and Group B—fetuses whose traces date were at least one week before labor.

	Group A (n = 27) Median (min-max), Mean ± SD or N (%)	Group B (n = 1045) Median (min-max), Mean ± SD or N (%)	P-Value
Baseline	134 (123–160)	137 (105–169)	0.507
Basal line	130 (122–146)	134 (105–168)	0.234
nAccel	5 (0–11)	5 (0–31)	0.714
nContr	1 (0–11)	1 (0–15)	0.246
mDec	0 (0–2)	0 (0–13)	0.175
iDec (% of no iDec)	22 (81.48)	1029 (98.47)	**<0.001**
pDec (% of no pDec)	27 (100)	1042 (99.71)	1.000
abSTV	52.89 ± 8.95	50.22 ± 8.44	0.105
avSTV	13.78 ± 3.65	14.57 ± 3.44	0.240
abLTV	3 (0–31)	0 (0–38)	**0.012**
avLTV	14.7 (8–33)	16.8 (0–40)	0.126
mean_UC	161.167 ± 138.37	167.342 ± 100.739	0.756
sd_UC	55.844 ± 44.593	46.480 ± 35.659	0.181
cv_UC	0.463 ± 0.329	0.372 ± 0.329	0.155
Gzip_UC	6.132 ± 1.981	5.691 ± 1.579	0.261
SampEn_UC	0.537 ± 0.269	0.592 ± 0.290	0.325
Gzip_FHR	11.356 ± 1.089	11.750 ± 0.883	**0.023**
SampEn_FHR	0.655 ± 0.149	0.692 ± 0.193	0.320

References

[1]. Lawn, J.E. Blencowe, H.; Oza, S.; You, D.; Lee, A.C.; Waiswa, P.; Lalli, M.; Bhutta, P.Z.; Barros, A.J.; Christian, P. Every newborn: Progress, priorities, and potential beyond survival. *Lancet* **2014**, *384*, 189–205.

[2]. Liu, L.; Oza, S.; Hogan, D.; Perin, J.; Rudan, P.; Lawn, P.J.E.L.; Cousens, P.S.; Mathers, C.; Robert, P. 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*, 430–440.

[3]. Blencowe, H.; Cousens, S.; Chou, D.; Oestergaard, M.; Say, L.; Moller, A.-N.; Kinney, M.; Lawn, J. Born too soon: The global epidemiology of 15 million preterm births. *Reprod. Health* **2013**, *10*, S2, doi:10.1186/1742-4755-10-S1-S2.

[4]. Lawn, J.E.; Gravett, M.G.; Nunes, T.M.; Rubens, C.E.; Stanton, C. Global report on preterm birth and stillbirth (1 of 7): Definitions, description of the burden and opportunities to improve data. *BMC Pregnancy Childbirth* **2010**, *10*, S1, doi:10.1186/1471-2398-10S1-S1.

[5]. Mwaniki, M.K.; Atieno, M.; Lawn, J.E.; Newton, C.R.J.C. Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: A systematic review. *Lancet* **2012**, *379*, 445–452.

[6]. Ayres-De-Campos, D.; Bernardes, J.; Costa-Pereira, A.; Pereira-Leite, L. Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. *Br. J. Obstet. Gynaecol.* **1999**, *106*, 1307–1310.

[7]. Bernardes, J.; Ayres-de-Campos, D. The persistent challenge of fetal heart rate monitoring. *Curr. Opin. Obstet. Gynecol.* **2010**, *22*, 104–109.

[8]. Spencer, J.A. Role of cardiotocography. *Br. J. Hosp. Med*. **1992**, *48*, 115–118.

[9]. Goncalves, H.; Pinto, P.; Ayres-de-Campos, D.; Bernardes, J. External uterine contractions signal analysis in relation to labor progression and dystocia. *IFMBE Proc.* **2014**, *41*, 555–558.

[10]. Goncalves, H.; Morais, M.; Pinto, P.; Ayres-de-Campos, D.; Bernardes, J. Linear and nonlinear analysis of uterine contraction signals obtained with tocodynamometry in prediction of operative vaginal delivery. *J. Perinat. Med.* **2017**, *45*, 327–332.

[11]. Bernardes, J.; Costa-Pereira, A.; Ayres-de-Campos, D.; van Geijn, H.P.; Pereira-Leite, L. Evaluation of interobserver agreement of cardiotocograms. *Int. J. Gynecol. Obstet.* **1997**, *57*, 33–37.

[12]. Donker, D.K.; Vangeijn, H.P.; Hasman, A. Interobserver variation in the assessment of fetal heart-rate recordings. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **1993**, *52*, 21–28.

[13]. Schiermeier, S.; von Steinburg, S.P.; Thieme, A.; Reinhard, J.; Daumer, M.; Scholz, M.; Hatzmann, W.; Schneider, K. Sensitivity and specificity of intrapartum computerised FIGO criteria for cardiotocography and fetal scalp pH during labor: Multicentre, observational study. *BJOG Int. J. Obstet. Gynaecol.* **2008**, *115*, 1557–1563.

[14]. Plsek, P.E.; Greenhalgh, T. Complexity science—The challenge of complexity in health care. *Br. Med. J.* **2001**, *323*, 625–628.

[15]. Parer, J.T. *Handbook of Fetal Heart Rate Monitoring*; W.B. Saunders Company: Philadelphia, PA, USA, 1997.

[16]. Reinl, E.L.; England, S.K. Fetal-to-maternal signaling to initiate parturition. *J Clin. Invest*. **2015**, *125*, 2569-2571.

[17]. Rotmensch, S.; Liberati, M.; Vishe, 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 Obstet. Gynecol. Scand.* **1999**, *78*, 493–500.

[18]. Signorini, M.G.; 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.

[19]. Nunes, I.; Ayres-de-Campos, D.; Figueiredo, C.; Bernardes, J. An overview of central fetal monitoring systems in labor. *J. Perinat. Med.* **2013**, *41*, 93–99.

[20]. Wilson, T.; Holt, T. Complexity science—Complexity and clinical care. *Br. Med. J.* **2001**, *323*, 685–688.

[21]. Chudacek, V.; Jiri, S.; Huptych, M.; Georgoulas, G.; Janku, P.; Koucky, M.; Stylios, C.; Lhotska, L. Automatic Classification of Intrapartal Fetal Heart-Rate Recordings—Can. It Compete with Experts? 2010. In *Proceedings of theInternational Conference on Information Technology in Bio-and Medical Informatics*; Springer: Berlin, Heidelberg, 2010; pp. 57–66.

[22]. Ayres-de Campos, D.; Berbardes, J.; Garrido, A.; Marques-de-sa, J.; Pereira-Leite, L. SisPorto 2.0: A program for automated analysis of cardiotocograms. *J. Matern. Fetal Med*. **2000**, *9*, 311–318.

[23]. Ayres-de-Campos, D.; Sousa, P.; Costa, A.; Bernardes, J. Omniview-SisPorto 3.5—a central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis. *J. Perinat. Med*. **2008**, *36*, 260–264.

[24]. 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*, 62–67.

[25]. 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.

[26]. Nunes, I.; Ayres-de Campos, D.; Austin, U.; Pina, A.; Philip, B.; Antony, N.; Simon, C.; Paulo, S.; Cristina, C.-S.; Joao, B. Central fetal monitoring with and without computer analysis a randomized controlled trial. *Obstet. Gynecol.* **2017**, *129*, 83–90.

[27]. Kolmogorov, A.N. Three approaches to the definition of the concept “quantity of information. *Probl. Peredachi Inf*. **1965**, *1*, 1–11.

[28]. Cilibrasi, R.; Vitanyi, P.; de Wolf, R. Algorithmic clustering of music based on string compression. *Comput. Music J.* **2004**, *28*, 49–67.

[29]. Cilibrasi, R.; Vitanyi, P.M.B. Clustering by compression. *IEEE Trans. Inf. Theory* **2005**, *51*, 1523–1545.

[30]. Wehner, S. Analyzing worms and network traffic using compression. *J. Comput. Secur.* **2005**, *15*, 303–320.

[31]. Deutsch, P. *DEFLATE Compressed Data Format Specification Version 1.3*; R.F.C. 1951, DOI 10.17487/RFC1951, Ed., 1996. https://www.rfc-editor.org/info/rfc1951

[32]. Ziv, J.; Lempel, A. Universal Algorithm for Sequential Data Compression. *IEEE Trans. Inf. Theory* **1977**, *23*, 337–343.

[33]. Huffman, D.A. A method for the construction of minimum-redundancy codes. *Proc. Inst. Radio Eng.* **1952**, *40*, 1098–1101.

[34]. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301.

[35]. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049.

[36]. Richman, J.; Moorman, R. Time series analysis using approximate entropy and sample entropy. *Biophys. J.* **2000**, *78*, 218a.

[37]. Goncalves, H.; Rocha, A.P.; 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 Hematol. Disord Drug Targets* **2008**, *8*, 91–98.

[38]. Amorim-Costa, C.; de Campos, D.A.; Bernardes, J. Cardiotocographic parameters in small-for-gestational-age fetuses: How do they vary from normal at different gestational ages? A study of 11687 fetuses from 25 to 40 weeks of pregnancy. *J. Obstet. Gynaecol. Res*. **2017**, *43*, 476–485.

[39]. Hewlett-Packard. *Hewlett-Packard Series 50 Service Manual.* 1993. M1351A and 1353A.

[40]. Goncalves, H.; ROCHA, A.P.; Ayres-de Campos, D.; Bernardes, J. Internal versus external intrapartum fetal heart rate monitoring: The effect on linear and nonlinear parameters. *Physiol. Meas*. **2006**, *27*, 307–319.

[41]. Monteiro-Santos, J.; Goncalves, 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.

[42]. Goncalves, H.; Costa, A.; Ayres-de-Campos, D.; Costa-Santos, C.; Rocha, A.P.; Bernaedes, J. Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of fetal heart rate variability. *Med. Biol. Eng. Comput*. **2013**, *51*, 665–676.

[43]. Lu, S.; Chen, X.; Kanters, J.K.; Solomon, I.C.; Chon, K.H. Automatic selection of the threshold value r for approximate entropy. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1966–1972.

[44]. Govindan, R.B.; Wilson, J.D.; Eswaran, H.; Lowery, C.L.; Peribl, H. Revisiting sample entropy analysis. *Phys. Stat. Mechan. Appl.* **2007**, *376*, 158–164.

[45]. Marzbanrad, F.; Kimura, Y.; Endo, M.; Palaniswami, M.; Khandoker, A.H. Transfer entropy analysis of maternal and fetal heart rate coupling. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2015**, *2015*, 7865–7868.

[46]. Moraes, E.R.; Murta, L.O.; Baffa, O.; Wakai, R.T.; Comani, S. Linear and nonlinear measures of fetal heart rate patterns evaluated on very short fetal magnetocardiograms. *Physiol. Meas*. **2012**, *33*, 1563–1583.

[47]. Padhye, N.S.; Brazdeikis, A.; Verklan, M.T. Change in complexity of fetal heart rate variability. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2006**, *1*, 1796–1798.

[48]. Van Leeuwen, P.; Cysarz, D.; Edelhauser, F.; Gronemeyer, D. Heart rate variability in the individual fetus. *Auton Neurosci.* **2013**, *178*, 24–28.

[49]. Bernardes, J.; Goncalves, H.; Ayres-de-Campos, D.; Rocha, A.P. 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*, 168–176.

[50]. Kim, K.N.; Park, Y.S.; Hoh, J.K. Sex-related differences in the development of fetal heart rate dynamics. *Early Hum. Dev.* **2016**, *93*, 47–55.

[51]. Ferrario, M.; Signorini, M.G.; Magenes, G. New indexes from the fetal heart rate analysis for the identification of severe intra uterine growth restricted fetuses. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2006**, *1*, 1458–1461.

[52]. Chourasia, V.S.T.; Anil, K. fetal heart rate variability analysis from phonocardiographic recordings. *J. Mechan. Med. Biol.* **2011**, *11*,1315–1331.

[53]. Pincus, S.M.; Viscarello, R.R. Approximate entropy: A regularity measure for fetal heart rate analysis. *Obstet. Gynecol.* **1992**, *79*, 249–255.

[54]. Hellgren, C.; Akerud, H.; Jonsson, M.; Poromaa, I.S. Sympathetic reactivity in late pregnancy is related to labor onset in women. *Stress Int. J. Biol. Stress* **2011**, *14*, 627-633.

ON THE PREDICTION OF FETAL ACIDAEMIA: A SPECTRAL ANALYSIS-BASED APPROACH

MOHAMMAD NOZARI, JOÃO SANTOS, LUISA CASTRO, CRISTINA SANTOS, JOÃO BERNARDES, ANTÓNIA COSTA

COMPUTERS IN BIOLOGY AND MEDICINE

VOLUME 109, PAGES 235-241

JUNE 2019

5. On the prediction of fetal acidemia: A spectral analysis-based approach

Abstract

A computational analysis of physiological systems has been used to support the understanding of how these systems work, and in the case of fetal heart rate, many different approaches have been developed in the last decades.

Our objective was to apply a new method of classification, which is based on spectral analysis, in fetal heart rate (FHR) traces to predict fetal acidosis diagnosed with umbilical arterial blood pH ≤ 7.05. Fast Fourier transform was applied to a real database for the classification approach. To evaluate the models, sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve were used.

Sensitivity equal to 1, specificity equal to 0.85 and an area under the ROC curve of 0.94 were found. In addition, when the definition of metabolic acidosis of umbilical arterial blood pH ≤ 7.05 and base excess ≤ −10 mmol/L was used, the proposed methodology obtained sensitivity = 1, specificity = 0.97 and area under the ROC curve = 0.98. The proposed methodology relies exclusively on the spectral frequency decomposition of the FHR signal. After further successful validation in more datasets, this approach can be incorporated easily in clinical practice due to its simple implementation. Likewise, the incorporation of this novel technique in an intrapartum monitoring station should be straightforward, thus enabling the assistance of labor professionals in the anticipated detection of acidaemia.

Introduction

Fetal heart rate (FHR) monitoring has been used for decades for the fetal well-being assessment before and during labor. Fetal acidaemia typically occurs when the fetus is deprived of oxygen during birth. Intrapartum fetal hypoxia can lead to serious consequences

for the newborn, specifically death and brain damage. Precocious obstetrical intervention depends on earlier diagnosis, and it is the cornerstone for fetal damage prevention. Umbilical arterial blood pH and base excess deficit in the extracellular fluid objectively reflect the fetal-newborn acid-base status and, consequently, its cellular oxygen supply. Cardiotocography (CTG), with its simultaneous evaluation of FHR and uterine contractions, is the most-used method for fetal well-being assessment, but its visual analysis has limitations with compromised reproducibility [1–3]. This is due to the highly complex nature of FHR traces. Plsek and Greenhalgh [4] conjectured that health components may be interpreted as simple (associated with high agreement and certainty), complex (associated with dynamic behaviors and interactions, thereby causing intermediate levels of agreement and certainty) or chaotic (no apparent visual information is retrieved leading to low levels of agreement and certainty). For this, FHR in labor is believed to vacillate between the complex and chaotic domains. To solve this, a computerized FHR analysis that uses signal processing and pattern recognition tools has been developed, which is supported by the idea that the analysis of FHR variability (FHRV) can provide additional information that is mainly related to the fetal autonomous nervous system (ANS) control of the heart [5]. Tools such as fractal [6,7], entropy [8,9] and wavelet [10,11] analyses have been used for FHR analysis in research since the early ‘90s but never in clinical practice. However, to date, no FHR analysis method has been capable of finding what exactly distinguishes the fetus in normal conditions from those with acidaemia. Another approach that is widely used is the analysis of traces in the frequency spectrum, using Fourier transforms [12]. Here, Siira et al. pointed out that spectral analysis provides a tool to quantify rather small changes in FHRV that may remain undetected if only the visual interpretation of FHR tracings is used [13]. In particular, it was mentioned [14] that Fourier analysis could yield new information about the fetal central nervous system and cardiac function. Spectral analysis is a method that can be used to objectively detect and quantify these changes in heart rate [15,16].

A wide variety of approaches have been considered using Fourier transformations to quantify the FHR variability by detecting QRS complexes [17] and to build methods that separate the fetal from maternal heart rate [18,19].

Chung et al. [20] stated that the use of real-time power spectral analysis could be a useful tool in diagnosing fetal well-being. They used Fourier transformation to investigate the power distributions among very low frequency (VLF), low frequency (LF), and high frequency bands to perform a linear analysis of the fetal heart rate. This study comprised 76 pregnant women, with the aim to predict fetal distress (abnormal FHR tracing and UAB pH < 7.15) and metabolic acidaemia (UAB pH < 7.15 and extracellular base excess deficit (BE) < -8 mmol/L). The results provide evidence that LF and VLF are good predictors of fetal distress (sensitivity 0.975, 0.75 and specificity 0.861, 0.944), while LF is a good predictor for fetal acidaemia (sensitivity 0.975 and specificity 0.861) [20]. In 2005, Maeda and Nagasawa [21] successfully applied fast Fourier transform to separate pathologic fetal sinusoidal heart rate traces from healthy ones, thus supporting the idea of the possibility for an automated real-time diagnostic tool. Rantonen et al. [22] found that fetuses with a UAB base deficit between 8 and 12 mmol/L showed decreased FHRV when assessed by spectral analysis.

In 2009, a study for neonatal acidaemia prediction (N = 148 FHR intrapartum tracings) was conducted using Omniview-SisPorto 3.5, which is a central fetal monitoring station with online alerts [23], based on combined computerized CTG and ST event linear analysis [24]. The prediction of cases with UAB pH ≤ 7.05 relying on red alerts from Omniview SisPorto using CTG and ST event analysis resulted in a sensitivity of 1 (95% CI, 0.560–1.00) and a specificity of 0.94 (95% CI, 0.89–0.97). However, when exclusively evaluating FHR signal (without ST data), the obtained sensitivity was 0.57 (95% CI, 0.20–0.88), and the specificity was 0.97 (95% CI, 0.92–0.99) [24].

Recently, in a comparison study of two classification systems based on linear indices, observers who were blind to clinical and outcome data classified tracings for severe metabolic acidaemia (UAB pH < 7.0 and BE < −12 mmol/L). Classification according to FIGO 3-tier system guidelines for pathological cases resulted in a sensitivity and specificity of 0.714 and 0.74, respectively, while the 5-tier system proposed by Parer and Ikeda, considering orange and red alerts, resulted in values of 0.619 and 0.801, respectively [25]. A hybrid approach combining linear and nonlinear methods achieved a sensitivity of 0.8 and a specificity of 0.71 with a dataset comprising 48 normal, 10 mildly acidemic and 10 moderate-to-severely

acidemic fetuses. This classification strategy used the approximate entropy combined with FHR's interval index [26].

Recently, the authors of this paper participated in a Challenge at the 2nd Signal Processing and Monitoring in Labor Workshop (SPaM) [27] using a novel classification methodology of fetal heart rate traces for acidemic prediction (pH < 7.05) using spectral analysis. The SPaM dataset comprises 300 intrapartum CTG tracings referring to three hospitals (Lyon, Oxford, Brno), which were provided by the CTG Challenge. All participants of the CTG Challenge were informed that 60 tracings referred to acidemic cases and had no other information regarding the outcome of newborns. Each team had five attempts to submit classification results, and the authors of this paper used the last three to tune the approach presented herein.

The CTG Challenge organizers provided the performance results of the approach presented by the authors for this paper, which achieved one of the best results among the classification methodologies in the contest. According to the CTG Challenge organization, this team obtained a score = sensitivity x specificity of 0.73, with a sensitivity and specificity of 0.67 and 0.80, respectively, which were known a posteriori.

In this work, we aimed to present our novel classification methodology of FHR traces for acidemic prediction using spectral analysis and applied it to another real dataset and tested it in an independent open-source dataset.

Methods

Data

The pre-existing database contains 148 FHR traces that were obtained for research purposes in a tertiary-care university hospital [24], together with biometric variables from the newborn and the mother, including information on the intrapartum fetal acid-base status (UAB pH and BE).

All FHR traces refer to singleton pregnancies that had more than 36 completed weeks were more than 60 min long and were sampled at 4 Hz. The interval between the tracing end and delivery was less than 5 and 20 min for vaginal and caesarean births, respectively.

Additionally, to test our approach in an independent dataset and to compare the results with those of other models, we used an open-access database with 552 traces from the Czech Technical University (CTU) in Prague and the University Hospital in Brno (UHB), CTU-UHB, as described in [28].

Signal Processing

Preprocessing of the signal was required to remove signal loss due to artefacts that are commonly present in the final minutes of FHR recordings immediately preceding the delivery. The preprocessing algorithm, which is described in detail in [29], detects samples lower than 60 bpm, samples above 200 bpm and consecutive differences higher than 25 bpm. The detected segments were then replaced by linear interpolation if referring to less than 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 50 min were selected for the following analyses.

Additionally, and only for the classic power spectral analysis indices computation, FHR signals were normalized, and the periodogram method was applied.

All these preprocessing computations were performed using MATLAB (R2018a, MathWorks, Natick, MA, USA).

Parameters and classification approach

A signal can be decomposed into its frequency components by using fast Fourier transform (FFT), which is an efficient algorithm that rapidly computes the discrete Fourier transformation [30]. For the 50 min of monitoring of each FHR in our datasets, an FFT was

computed by using R language [31], and the magnitude of each complex number (peak) was calculated. For the decision feature, two arguments were defined: an amplitude cut-off (c) and the number of peaks above that cut-off (p-number).

For the FFT of each FHR signal, the minimum and maximum frequency amplitude values were obtained, and the global minimum and maximum were computed across all signals. Then, for each p-number from 1 to 10 and through the Newton method of successive approximations [32], the minimum amplitude cut-off, which classifies all signals as abnormal, is set as minc. Likewise, the maximum amplitude that classifies all signals as normal is set as maxc. For each signal and for each p-number, we let the amplitude cut-off slide between minc and max c by the steps of (maxc-minc)/50, thus predicting the classification for each (p-number, c) pair.

For an FHR signal, if its FFT has at least p-number peaks with an amplitude above threshold c, then the FHR recording is classified as abnormal; otherwise, it is labelled as normal (see Fig. 5.1 for an example of each). For each p-number between 1 and 10, the performance metrics are computed for the set of 50 amplitude cut-offs.

To compare the approach presented in this paper with the classic indices that were derived from power spectral analysis (PSA), we also performed the integration of the spectrum over specific frequency bands. The appropriate bandwidths for the FHR signals analysis [5] were used: very low frequency (VLF: 0–0.03 Hz), low frequency (LF: 0.03–0.15 Hz), high frequency (HF: 0.5–1 Hz) and movement frequency (MF: 0.15–0.5 Hz).

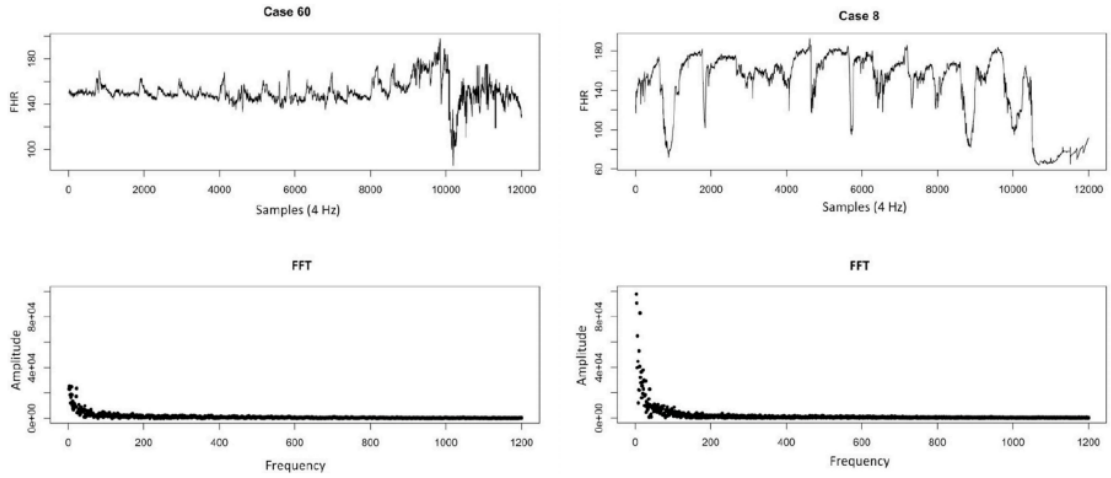

Figure 5.1. Fetal heart rate sampled at 4 Hz (50 min) for an example of a non-acidemic fetus (top left) and an acidemic one (top right). The corresponding spectra with frequency amplitudes are presented below for non-acidemic (left) and for the acidemic (right).

Evaluation metrics and outcome

The classifications of the 148 FHR traces, using an FFT approach, were evaluated by comparing the classification results with the following fetal acidaemia definition: UAB pH ≤ 7.05 for abnormal traces (fetal acidaemia). The same approach was also evaluated considering the base excess in the definition of acidemic fetuses (UAB pH ≤ 7.05 and BE ≤ −10 mmol/L for abnormal traces).

The classification performance was assessed by the means of sensitivity, specificity and an area under the ROC curve. Sensitivity measures the ability of the classifier to correctly identify acidemic fetuses, while the specificity evaluates the ability of the classifier to identify non-acidemic fetuses. For each pnumber, the ROC curve was obtained relating the specificity and sensitivity of the classifier for each cut-off.

Results

As in the CTG challenge, in our dataset, the objective is to classify the 148 FHR recordings as acidemic according to the same pH cut-off of 7.05. From the 148 cases, only 7 referred to newborns' arterial umbilical blood with pH ≤ 7.05.

For each p-number value, the area under the ROC curve was obtained (Table 5.1), having achieved maximum values of 0.938 and 0.939 for p-numbers in the range of 3–4 peaks above the threshold (Fig. 5.2 and Table 5.1). The highest score, sensitivity x specificity, was 0.922 (Table 2) and pertained to the case of 4 peaks above the 27th step in the cut-off parameter and corresponded to a sensitivity of 1 (95% CI 0.561–1.000) and a specificity of 0.851 (95% CI 0.779–0.903).

Table 5.1. Area under ROC curve computed for each value of parameter p-number and 50 steps for the cut-off parameter, for the prediction of pH≤7.05.

p-number	1	2	3	4	5	6	7	8	9	10
AUC	0.831	0.907	0.938	0.939	0.860	0.887	0.895	0.879	0.895	0.893

Table 5.2. Performance metrics with selected parameters for the prediction of pH≤7.05. Sensitivity, specificity and score values with respective confidence interval (CI) for p-number between 3 and 4 and steps with the highest score values (>0,8) for some p-numbers.

amplitude cut-off	p-number = 3			p-number = 4		
c	sensitivity [95% CI]	specificity [95% CI]	score	sensitivity [95% CI]	specificity [95% CI]	score
20	0.714 [0.303, 0.949]	0.957 [0.906, 0.983]	0.827	0.429 [0.118, 0.798]	0.957 [0.906, 0.983]	0.641
21	0.714 [0.303, 0.949]	0.915 [0.853, 0.953]	0.808	0.571 [0.202, 0.882]	0.957 [0.906, 0.983]	0.739
22	0.714 [0.303, 0.949]	0.901 [0.836, 0.943]	0.802	0.571 [0.202, 0.882]	0.943 [0.888, 0.973]	0.734
23	0.714 [0.303, 0.949]	0.887 [0.819, 0.932]	0.796	0.714 [0.303, 0.949]	0.929 [0.870, 0.964]	0.814
24	0.714 [0.303, 0.949]	0.887 [0.819, 0.932]	0.796	0.714 [0.303, 0.949]	0.915 [0.853, 0.953]	0.808
25	0.857 [0.420, 0.992]	0.879 [0.811, 0.926]	0.868	0.714 [0.303, 0.949]	0.879 [0.811, 0.926]	0.792
26	0.857 [0.420, 0.992]	0.844 [0.771, 0.898]	0.850	0.857 [0.420, 0.992]	0.865 [0.795, 0.915]	0.861
27	0.857 [0.420, 0.992]	0.823 [0.747, 0.880]	0.840	**1.000 [0.561, 1.000]**	**0.851 [0.779, 0.903]**	**0.922**
28	0.857 [0.420, 0.992]	0.809 [0.732, 0.868]	0.833	1.000 [0.561, 1.000]	0.801 [0.724, 0.862]	0.895
29	1.000 [0.561, 1.000]	0.787 [0.709, 0.850]	0.887	1.000 [0.561, 1.000]	0.752 [0.671, 0.819]	0.867
30	1.000 [0.561, 1.000]	0.766 [0.686, 0.831]	0.875	1.000 [0.561, 1.000]	0.738 [0.656, 0.806]	0.859
31	1.000 [0.561, 1.000]	0.709 [0.626, 0.781]	0.842	1.000 [0.561, 1.000]	0.688 [0.604, 0.762]	0.829
32	1.000 [0.561, 1.000]	0.695 [0.611, 0.768]	0.834	1.000 [0.561, 1.000]	0.652 [0.567, 0.729]	0.807

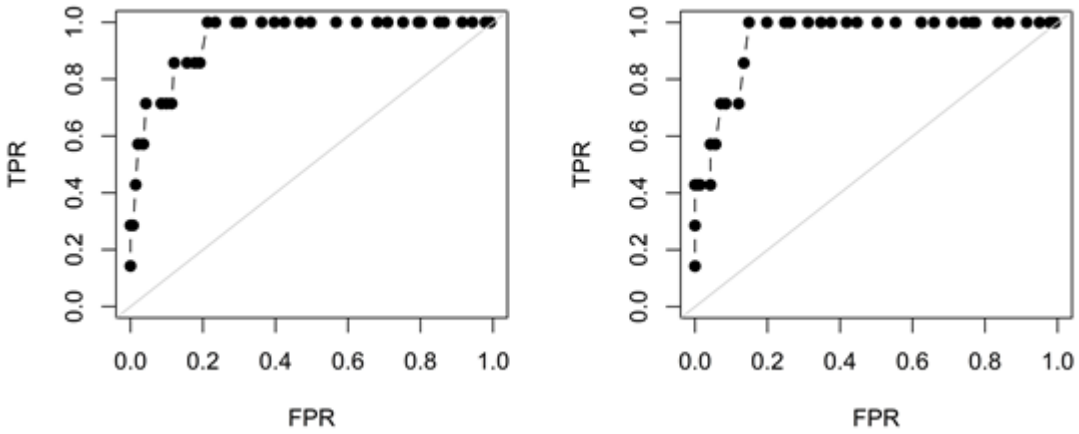

Figure 5.2. ROC curves corresponding to the two highest areas for the prediction of newborns with pH≤7.05: 0.938 for p-number=3 (left) and 0.939 for p-number=4 (right).

Some of the state-of-the-art methodologies also include the umbilical arterial blood (UAB) base excess to define neonate metabolic acidaemia, even though its utility has been questioned [33]. Hence, and for comparison purposes, we also performed the classification of the dataset according to this outcome, defining the acidemic group as FHR traces corresponding to newborns with pH ≤ 7.05 and a base excess (BE)≤-10 mmol/L. From the 148 cases, only 6 referred to newborns with UAB pH ≤ 7.05 and BE ≤ −10 mmol/L. The p-numbers resulting in a higher area under the ROC curve were in the range of 6–7 peaks above the cut-off, with 0.975 and 0.980 (Table 5.3). The highest score for the dataset, in the prediction of UAB pH ≤ 7.05 and BE ≤ −10 mmol/L, was 0.986 (Table 5.4), referring to the case of 7 peaks above the 21st step in the cut-off parameter, corresponding to a sensitivity of 1 (95% CI 0.517–1.000) and a specificity of 0.972 (95% CI 0.925–0.991).

Table 5.3. Area under the curve (AUC) computed for each value of parameter p-number and 50 steps for the cut-off parameter for the prediction of newborns with UAB pH≤7.05 and BE≤-10 mmol/L.

p-number	1	2	3	4	5	6	7	8	9	10
AUC	0.802	0.889	0.923	0.939	0.954	0.975	0.980	0.960	0.967	0.961

For comparison reasons, the integration of the FFT spectrum over the specific frequency bands commonly used in FHR signals [5] was also assessed. Indices from the classic HRV spectral analysis did not achieve the performance metrics from the simple and original approach that was presented in this paper (Table 5.5).

Table 5.4. Performance metrics with selected parameters for the prediction of newborns with UAB pH≤7.05 and BE≤-10 mmol/L. Sensitivity, specificity and score values for p-number between 6 and 7 and steps with the highest score values (>0,8) for some p-numbers.

amplitude cut-off	p-number = 6			p-number = 7		
c	sensitivity [95% CI]	specificity [95% CI]	score	sensitivity [95% CI]	specificity [95% CI]	score
19	0.500 [0.188, 0.812]	0.993 [0.956, 1.000]	0.705	0.667 [0.241, 0.940]	0.979 [0.935, 0.995]	0.808
20	0.500 [0.188, 0.812]	0.979 [0.935, 0.995]	0.700	0.667 [0.241, 0.940]	0.972 [0.925, 0.991]	0.805
21	0.500 [0.188, 0.812]	0.972 [0.925, 0.991]	0.697	1.000 [0.517, 1.000]	0.972 [0.925, 0.991]	0.986
22	0.833 [0.365, 0.991]	0.965 [0.916, 0.987]	0.897	1.000 [0.517, 1.000]	0.958 [0.906, 0.983]	0.979
23	0.833 [0.365, 0.991]	0.965 [0.916, 0.987]	0.897	1.000 [0.517, 1.000]	0.930 [0.871, 0.964]	0.964
24	1.000 [0.517, 1.000]	0.944 [0.888, 0.974]	0.972	1.000 [0.517, 1.000]	0.901 [0.837, 0.943]	0.949
25	1.000 [0.517, 1.000]	0.915 [0.854, 0.954]	0.957	1.000 [0.517, 1.000]	0.880 [0.813, 0.927]	0.938
26	1.000 [0.517, 1.000]	0.894 [0.829, 0.938]	0.946	1.000 [0.517, 1.000]	0.852 [0.781, 0.904]	0.923
27	1.000 [0.517, 1.000]	0.859 [0.788, 0.910]	0.927	1.000 [0.517, 1.000]	0.845 [0.773, 0.898]	0.919
28	1.000 [0.517, 1.000]	0.831 [0.757, 0.887]	0.912	1.000 [0.517, 1.000]	0.810 [0.734, 0.869]	0.900
29	1.000 [0.517, 1.000]	0.782 [0.703, 0.845]	0.884	1.000 [0.517, 1.000]	0.775 [0.695, 0.839]	0.880
30	1.000 [0.517, 1.000]	0.754 [0.673, 0.820]	0.868	1.000 [0.517, 1.000]	0.725 [0.643, 0.795]	0.851
31	1.000 [0.517, 1.000]	0.718 [0.636, 0.789]	0.847	1.000 [0.517, 1.000]	0.697 [0.614, 0.770]	0.835
32	1.000 [0.517, 1.000]	0.676 [0.592, 0.751]	0.822	1.000 [0.517, 1.000]	0.676 [0.592, 0.751]	0.822
33	1.000 [0.517, 1.000]	0.641 [0.556, 0.718]	0.801	1.000 [0.517, 1.000]	0.641 [0.556, 0.718]	0.801

Table 5.5. Area under the curve (AUC) (and respective confidence intervals) computed for each spectral band for (a) the prediction of pH≤7.05 and (b) for the prediction of pH≤7.05 and BE≤-10 mmol/L**.**

Spectral Bands	VLF	LF	HF	MF
AUC (a) [95% CI]	0.804* [0.687, 0.922]	0.819* [0.706, 0.931]	0.687 [0.480, 0.894]	0.710 [0.534, 886]
AUC (b) [95% CI]	0.772* [0.648, 0.896]	0.788* [0.669, 0.906]	0.657 [0.425, 0.890]	0.678 [0.484, 0.873]

***p<0.05**

To test the proposed approach in an independent set, the CTU-UHB open-access CTG intrapartum database was used [28]. After the pre-processing step, all 533 traces with respective UAB pH information were used for testing.

The amplitude range of the traces in the CTU-UHB differs substantially between the dataset in the CTG challenge and the dataset used for training our method. Since our methodology relies on the maximum amplitude of the traces, we selected the cut-off of 0.0185 of the maximum amplitude for the Physionet dataset, which corresponds to the mean of the ratio between the cut-off and the maximum amplitude used in the CTG challenge dataset, 0.0175, and in this paper's dataset, 0.0195. Thus, for the CTU-UHB, traces with more than 4 peaks above the cut-off of 0.0185 of the maximum amplitude were classified as acidemic. With this approach, we obtained results of 0.636 for sensitivity and 0.801 for specificity.

Discussion

In this work, a new classification methodology based on fast Fourier transformation of FHR signals is proposed. In this framework, an amplitude threshold and the number of frequencies (p-number) above that threshold are the only parameters to adjust. This approach was first applied in the context of the SPaM Challenge [27], which resulted in one of the best scores within the participant classifications. As such, another dataset was submitted to the methodology proposed herein, which aims for the replication of results. For both datasets, high scores were obtained. Moreover, the optimized score using only the pH feature resulted in a similar parameter p-number for both datasets, which suggested no data dependency.

Several studies have evaluated the accuracy of linear, nonlinear and hybrid approaches in the prediction of neonatal outcomes. The results that were obtained with this methodology from two real intrapartum FHR datasets compare with the performance metrics of state-of-the-art approaches. In the present study, for a p-number equal to 4 and cut-off in the 27th step, sensitivity and specificity were 1 (95% CI 0.561–1.000) and 0.851 (95% CI 0.779–0.903), respectively. This means that, in comparison with the previous study on the same dataset,

the proposed approach has an increased capacity of correctly identifying acidemic fetuses, which thus reduces the false negatives that were detected [24] (Table 5.6).

Table 5.6. Comparison of results of the present approach with the literature results.

Study	Dataset (size)	Outcome (acidemic)	Sensitivity	Specificity
Chung et al. [20]	Private (76)	pH < 7.15 and BE < -8	0.975	0.861
Costa, A et al. [24]	Private (148)	pH ≤ 7.05	0.570	0.970
Martí Gamboa et al. [25]	Private (202)	pH ≤ 7.0 and BE < -12 Method 1	0.714	0.740
Martí Gamboa et al. [25]	Private (202)	pH ≤ 7.0 and BE < -12 Method 2	0.619	0.801
Gonçalves, H. et al. [26]	Private (68)	pH ≤ 7.10	0.800	0.710
Authors' approach	Training Dataset (148)	pH ≤ 7.05	1.000	0.851
Rotariu et al. [34]	CTU-UHB	pH < 7.20 and BDef > 8	0.960	0.876
Rotariu et al. [35]	CTU-UHB	pH < 7.20	0.732	0.882
Cömert [36]	CTU-UHB	pH < 7.20	0.778	0.768
Pasarica [37]	CTU-UHB	pH < 7.15	0.938	0.831
Authors' approach	CTU-UHB	pH ≤ 7.05	0.636	0.801

More generally, the results that were obtained in this paper outperform the reported performance metrics from other published methodologies for newborn acidemic prediction. The large confidence interval that was obtained for sensitivity is due to a short set of acidemic fetuses. However, the small confidence interval for specificity supports the evidence that this approach represents a promising strategy, although it needs validation in other datasets.

The results that were obtained in the test phase, with the independent open source database CTU-UHB, are promising and suggest further research to refine the cut-off selection strategy and to validate it in independent and larger databases.

These results might seem shy compared to those of other classification studies [34–37] that used the same test dataset, CTU-UHB (see Table 5.6). However, three have used a pH threshold of 7.20 [34–36] or 7.15 [37], while one also reports the testing performance of a small proportion of the original dataset [34]. In [38], the same 7.05 pH threshold was defined, and although a deep learning technique was applied in the training set of more than 30000 CTGs, only a small improvement was accomplished when it was applied to CTU-UHB, compared to using our simple approach.

Our results for two datasets of 300 and 148 cases and for the 552 cases of the testing set CTU-UHB (Table 5.7) support the evidence that our FFT approach to a frequency analysis of FHR signals is a promising strategy for the prediction of acidemic newborns that can outperform previous power spectral analysis (PSA) [39]. In PSA, we explore the analysis of FHR frequency bandwidths validated in animal and human studies, which were adapted to human fetuses. The total power reflects the overall activity of the physiological systems underlying the FHR control, specifically the sympathetic and parasympathetic (vagal) systems, which are associated with low frequency (LF) and high (HF) frequency bands, respectively, as well as the thermoregulatory and renin-angiotensin control systems that are associated with the VLF band [39]. In our FFT-based frequency analysis, we found an association between newborn acidaemia and four to seven FHR frequency peaks above a certain threshold of amplitude, which were determined by an algorithm that was specially developed by us (see Fig. 1). Peaks of FHR spectral frequencies, above a certain threshold, were consistent with the overall activation of the physiological FHR control systems, which were associated with the installation of fetal acidaemia during labor, were observed in PSA and comprised the activation of several physiological control components, such as the sympathetic and parasympathetic (vagal) systems, as well as the thermoregulatory and renin-angiotensin systems [39]. However, further studies are necessary to better elucidate the pathophysiological background of our findings and to confirm or refute it, as well as explain why our FFT-based analysis of FHR recordings performs better or worse than PSA, in the prediction of newborn acidaemia.

Table 5.7. Sensitivity and specificity obtained for the parameter selection with the highest score (of the two training datasets) and respective confidence intervals for the datasets that were analysed in this paper.

Dataset	Acidemic class rule	Sensitivity [95% CI]	Specificity [95% CI]
CTG Challenge	pH<7.05	0.67*	0.80*
dataset analysed	pH≤7.05	1.000 [0.561, 1.000]	0.851 [0.779, 0.903]
dataset analysed	pH≤7.05 and BE≤-10	1.000 [0.517, 1.000]	0.972 [0.925, 0.991]
CTU-UHB (test dataset)	pH≤7.05	0.636 [0.477, 0.772]	0.801 [0.763, 0.834]

* results provided by SPaM organizers

Focusing on the amplitude peaks of lower frequencies, our approach could represent an alternative methodology that supports previous studies, which refer to LF and VLF as good predictors for fetal acidaemia and distress classification [20]. The proposed methodology relies exclusively on the spectral frequency decomposition of the FHR signal being easily incorporated in an intrapartum monitoring station, thereby assisting labor professionals in the anticipated detection of acidaemia.

One limitation of this approach is that for large datasets, the parameter tuning is time-consuming. However, it would only be a one-time computation.

In future work, the use of smaller steps around the cut-off to find the global maximum for the score will be investigated. Different datasets will also be analyzed to validate, refine the cut-off selection strategy and better understand the data independency that was suggested by the current results.

Summary Points

What was already known about the topic:
• Precocious obstetrical intervention depends on earlier diagnosis, and it is the cornerstone of fetal damage prevention;
• Several linear and/or nonlinear methods for fetal heart rate analysis have been proposed for the prediction of fetal acidaemia with reasonable results, which involves a wide range of data analysis experts;
• In the literature, there is evidence that a traditional power
spectral analysis of fetal heart rate achieves high-performance results, thereby electing LF and VLF bands as good predictors of fetal acidaemia.

What this study added to our knowledge:

• The new spectral analysis approach described in this study provides further support that the power spectral analysis of FHR signals is a promising strategy for the prediction of acidemic newborns;
• The new approach to power spectral analysis that is presented here has promising classification performance in three real databases, which provides evidence that our model is data independent.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

This work was supported by the North Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund through the project "NORTE-01-0145-FEDER-000016" (NanoSTIMA). L. Castro acknowledges the Postdoc Grant with reference NanoSTIMA-WP4.1.1-D. J. Santos acknowledges the support of the doctorate scholarship from the Clinical Research and Health Services Program, funded by the European Social Fund under the Portuguese Norte 2020 research and innovation program (grant agreement n° Norte-08-5369-FSE-000063).

References

[1] D. Ayres-de-Campos, J. Bernardes, A. Costa-Pereira, L. Pereira-Leite, Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision, Br. J. Obstet. Gynaecol. 106 (1999) 1307–1310.

[2] J. Bernardes, D. Ayres-de-Campos, The persistent challenge of fetal heart rate monitoring, Curr. Opin. Obstet. Gynecol. 22 (2010) 104–109.

[3] J. Bernardes, A. Costa-Pereira, D. Ayres-de-Campos, H.P. van Geijn, L. Pereira-Leite, Evaluation of interobserver agreement of cardiotocograms, Int. J. Gynaecol. Obstet. 57 (1997) 33–37.

[4] P.E. Plsek, T. Greenhalgh, Complexity science: the challenge of complexity in health care, BMJ 323 (2001) 625–628.

[5] M.G. Signorini, G. Magenes, S. Cerutti, D. Arduini, Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings, IEEE Trans. Biomed. Eng. 50 (2003) 365–374.

[6] N.A. Gough, Fractals, chaos, and fetal heart rate, Lancet 339 (1992) 182–183.

[7] M.K. Yum, E.Y. Park, C.R. Kim, J.H. Hwang, Alterations in irregular and fractal heart rate behavior in growth-restricted fetuses, Eur. J. Obstet. Gynecol. Reprod. Biol. 94 (2001) 51–58.

[8] S.M. Pincus, Approximate entropy as a measure of system complexity, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 2297–2301.

[9] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. Heart Circ. Physiol. 278 (2000) H2039–H2049.

[10] S. Papadimitriou, D. Gatzounas, V. Papadopoulos, V. Tzigounis, A. Bezerianos, Denoising of the fetal heart rate signal with nonlinear filtering of the wavelet transform maxima, Int. J. Med. Inform. 44 (1997) 177–192.

[11] E. Salamalekis, P. Thomopoulos, D. Giannaris, I. Salloum, G. Vasios, A. Prentza, D. Koutsouris, Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilizing wavelet analysis and neural networks, BJOG 109 (2002) 1137–1142.

[12] H. Shono, M. Yamasaki, M. Muro, M. Oga, Y. Ito, K. Shimomura, H. Sugimori, Chaos, and fractals which 1/f spectrum below 10(-2) Hz demonstrates full-term fetal heart rate changes during the active phase, Early Hum. Dev. 27 (1991) 111–117.

[13] S.M. Siira, T.H. Ojala, T.J. Vahlberg, J.O. Jalonen, I.A. Välimäki, K.G. Rosén, E.M. Ekholm, Marked fetal acidosis and specific changes in power spectrum analysis of fetal heart rate variability recorded during the last hour of labor, BJOG 112 (2005) 418–423.

[14] D. Rassi, M.J. Lewis, Power spectral analysis of the fetal magnetocardiogram, Physiol. Meas. 16 (1995) 111–120.

[15] S. Akselrod, D. Gordon, F.A. Ubel, D.C. Shannon, A.C. Berger, R.J. Cohen, Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control, Science 213 (1981) 220–222.

[16] S. Siira, T. Ojala, E. Ekholm, T. Vahlberg, S. Blad, K.G. Rosén, Change in heart rate variability in relation to significant st-event associates with newborn metabolic acidosis, BJOG 114 (2007) 819–823.

[17] S. Abboud, D. Sadeh, Power spectrum analysis of fetal heart rate variability using the abdominal maternal electrocardiogram, J. Biomed. Eng. 12 (1990) 161–164.

[18] M. Haritopoulos, C. Capdessus, A.K. Nandi, Fetal pqrst extraction from ECG recordings using cyclostationarity-based source separation method, Conf Proc IEEE Eng Med Biol Soc (2010) 1910–1913 2010.

[19] C. Knapp, G. Carter, Generalized correlation method for estimation of time-delay, IEEE Trans. Acoust. Speech Signal Process. 24 (1976) 320–327.

[20] D.Y. Chung, Y.B. Sim, K.T. Park, S.H. Yi, J.C. Shin, S.P. Kim, Spectral analysis of fetal heart rate variability as a predictor of intrapartum fetal distress, Int. J. Gynaecol. Obstet. 73 (2001) 109–116.

[21] K. Maeda, T. Nagasawa, Automatic computerized diagnosis of fetal sinusoidal heart rate, Fetal Diagn. Ther. 20 (2005) 328–334.

[22] T. Rantonen, E. Ekholm, S. Siira, T. Metsälä, R. Leino, U. Ekblad, I. Välimäki, Periodic spectral components of fetal heart rate variability reflect the changes in cord arterial base deficit values: a preliminary report, Early Hum. Dev. 60 (2001) 233–238.

[23] D. Ayres-de-Campos, P. Sousa, A. Costa, J. Bernardes, Omniview-sisporto 3.5 – a central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis, J. Perinat. Med. 36 (2008) 260–264.

[24] A. Costa, D. Ayres-de-Campos, F. Costa, C. Santos, J. Bernardes, Prediction of neonatal acidemia by computer analysis of fetal heart rate and ST event signals, Am. J. Obstet. Gynecol. 201 (464) (2009) e461–466.

[25] S. Martí Gamboa, O.R. Giménez, J.P. Mancho, M.L. Moros, J.R. Sada, S.C. Mateo, Diagnostic accuracy of the figo and the 5-tier fetal heart rate classification systems in the detection of neonatal acidemia, Am. J. Perinatol. 34 (2017) 508–514.

[26] H. Gonçalves, A.P. Rocha, D. Ayres-de-Campos, J. Bernardes, Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery, Med. Biol. Eng. Comput. 44 (2006) 847–855.

[27] CTG Challenge at the 2nd Signal Processing and Monitoring (SPaM) in Labor Workshop, Oxford, United Kingdom, (30 October - 1 November 2017) [online] Available at: http://users.ox.ac.uk/~ndog0178/CTGchallenge2017.htm , Accessed date: 12 April 2018.

[28] V. Chudacek, J. Spilka, M. Bursa, P. Janku, L. Hruban, M. Huptych, et al., Open access intrapartum CTG database, BMC Pregnancy Childbirth 14 (2014) 16, https://doi.org/10.1186/1471-2393-14-16.

[29] H. Gonçalves, A.P. Rocha, D. Ayres-de-Campos, J. Bernardes, Internal versus external intrapartum fetal heart rate monitoring: the effect on linear and nonlinear parameters, Physiol. Meas. 27 (2006) 307–319.

[30] N. Ahmed, K.R. Rao, Fast fourier transform, Orthogonal Transforms for Digital Signal Processing, Springer, Berlin, Heidelberg, 1975.

[31] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2018 URL https://www.Rproject.org/.

[32] K. Atkinson, An Introduction to Numerical Analysis, Wiley, 1989, p. 712.

[33] L. Knutzen, E. Svirko, L. Impey, The significance of base deficit in acidemic term neonates, Am. J. Obstet. Gynecol. 213 (373) (2015) e371–377.

[34] Cristian Rotariu, Alexandru Pasarica, Hariton Costin, Dragos Nemescu, Spectral analysis of fetal heart rate variability associated with fetal acidosis and base deficit values, International Conference on Development and Application Systems (DAS), IEEE, 2014.

[35] Cristian Rotariu, Alexandru Pasarica, Gladiola Andruseac, Hariton Costin, Automatic analysis of the fetal heart rate variability and uterine contractions, International Conference and Exposition on Electrical and Power Engineering (EPE 2014), 2014.

[36] Zafer Cömert, Adnan FatihKocamaz, Open-access software for analysis of fetal heart rate signals, Biomed. Signal Process. Control 45 (August 2018) 98–108.

[37] Alexandru Pasarica, Cristian Rotariu, Radu Gabriel, Bozomitu, Oana Diana Eva, Dynamic of Couplings between Fetal Heart Rate and Uterine Contractions, IEEE, 2015.

[38] A. Petrozziello, I. Jordanov, T. Aris Papageorghiou, W.G. Christopher Redman, A. Georgieva, Deep Learning for Continuous Electronic Fetal Monitoring in Labor. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (2018).

[39] H. Gonçalves, A.P. Rocha, D. Ayres-de-Campos, J. Bernardes, Frequency domain and entropy analysis of fetal heart rate: appealing tools for fetal surveillance and pharmacodynamic assessment of drugs, Cardiovasc. Haematol. Disord. – Drug Targets 8 (2008) 91–98.

GENERAL DISCUSSION AND CONCLUSIONS

6. General Discussion and Conclusions

An efficient assessment of fetus and mother well-being during pregnancy and labor has required different generations of researchers much effort. Although the fetus's heart rate is its main welfare indicator, a lot is still unknown due to its complexity. The challenges include appropriate clinical decision support for CTG and satisfactory signal processing techniques [2].

Soon after the first steps of HRV analysis, it was shown that linear indexes were not capturing all the information contained in the signal. Therefore, more robust approaches started to emerge. The purpose of this thesis was to provide one more input to the researching community on how complexity analysis is a key factor in understanding the mechanisms underlying fHRV better.

There is no doubt of the importance of nonlinear measures in fetal monitoring, as they enrich the signal description by providing new indicators for classification and diagnostic purposes. In 2012, Fucher et al. [3] studied over 9000 time-series analysis features, and among those most correlated with umbilical cord pH were nonlinear measures. In a study by Choi *et al.* [4], nonlinear dynamic indices were able to differentiate normal pregnancies from ones with partial placental abruption with high accuracy, while linear indices were not. Hence, the integration of advanced signal processing approaches, linear and frequency domain indexes as a multiparametric approach, provide an improvement in the statistical analysis of biological signals, as shown by Spilka [5], Magenes [6], and Tetschke [7].

In daily clinical practice, the application of procedures relies heavily on the clinician's understanding of the underlying methodologies. Therefore, a physiological interpretation of complexity indexes and their association is of great importance. Consequently, this dissertation aims to study how do complexity measures relate to standard physiological characteristics of FRH tracings. In **Chapter 3,** we concluded that different information from fHRV is captured by the two nonlinear measures used, which shows their complementarity. It was also shown that complexity through different scales (multiscale) provides a clearer difference between groups of acidemic and non-acidemic babies [8]. In fact, another objective of this dissertation is to assess how can complexity measures or other different and new

approaches can predict fetal acidemia. On this matter, we created a spectral based algorithm for acidemia prediction during intrapartum, described in **Chapter 5**. The results of this new algorithm were compared with other approaches [9]. However, our methodology primes for its simplicity, which is a major advantage as the interpretability need in clinical practice is one of the constraints in implementing new approaches.

On a different topic, but also related to acidemia, is the question of when labor will take place. Another objective of this dissertation is to assess how well complexity measures can predict labor. In particular, acknowledging that a particular fetus is at risk of becoming premature, born before 37 weeks of gestation, would provide health care professionals information, and most important, time to redirect resources to this particular case [10]. In **Chapter 4**, we presented results suggesting that complexity analysis of FHRV, particularly using compression, can help predict how close to labor the fetus is.

Strengths and weaknesses

This dissertation has some weaknesses, like the ones related to the quality of recordings and the intrinsic subject variability and complexity of pathologies, complicating their prediction and control. Even with the definition of acidosis, there is some controversy, as different authors consider different pH cutoffs, and some include base excess or base deficit [11]. Some defined as "at risk of acidemia" when $pH < 7.20$ [8, 12] or $pH < 7.15$ [13-15]; others define when $pH < 7.1$ [16-19] or even when $pH < 7.05$ [5, 20-24]. Another challenge is to collect enough data for a proper acidemia analysis since the prevalence of an acidemic fetus ranges from 0.6% to 3.5% [25, 26]. The effect of an antepartum vs. intrapartum analysis on the complexity indices and the differences the signal acquisition methods produce are also important variables to take into account to correctly evaluate and assess fetus well-being [27, 28]. Much effort has been put into the signal acquisition and processing models because the extracted features highly rely on the quality of the preprocessing steps, such as artifacts removal, interpolation method, segmentation, and detrending of the signal [29]. FHRV analysis depends on its non-stationarity (properties like mean, variance, and correlation structure vary in time through events like uterine contractions). One way to counter this is to

select small temporal windows where this property holds. It is considered an interval of 10-20 min for the minimum time windows to perform the analysis for tracing classification and clinical decision [30, 31].

FHR monitors acquire the beat-to-beat intervals in milliseconds either from Doppler or electrocardiographic signals and then converts them to provide a sequence of instantaneous heart rates, in beats per minute (bpm). However, when data is exported, it has to be sampled, meaning an interpolation of signals [32]. The sampling rate does not seem to affect many linear parameters, but when nonlinear ones are considered, differences were found [28]. This is a crucial issue when defining reference values for irregularity indices, such as Entropy, as they depend on the sampling frequency, as shown in [28], where 2Hz *vs.* 4Hz sampling was compared. Thus, it is essential not to compare computerized fHR analysis systems that use different sampling rates [30, 33].

This dissertation has the strength of relating linear and nonlinear indexes. Understanding the relation between linear and nonlinear indexes makes it easier to interpret nonlinear indices and develop new methods for predicting poor neonatal outcomes.

Future work

Nonlinear methodologies must continue to be studied and applied to retrieve signal with the best quality possible, dismissing as much noise as possible. This is important when adopting low-cost systems for signal extraction, as is the case of the fetal phonocardiography, which has a low signal-to-noise ratio [34].

Usage of continuous noninvasive evaluation, such as the usage of wearables, has been discussed [35, 36] and will contribute to the patient 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' development, so approaches taking into consideration both linear and nonlinear measures, through multivariate analysis, can improve the assessment of fetal well-being, and consequently,

maternal well-being. The association between fetal heart rate and uterine contractions, and their relationship regarding fetal response to the rigid environment it is subject to during labor (mainly), is still an overlooked research area. Future studies should be focused since the fetus has to constantly adapt to its environment. The same can be said regarding the relation of fHR with the maternal heart rate. Different analysis methods take into account the maternal heart rate but mainly as a noise reduction factor to better extract fetal heart rate. Instead, it should be seen as a crucial variable as it might reflect both organisms' interaction [37, 38].

Conclusion

This work highlights the complexity of this field of study. The scientific community needs to keep making steps so that fetal mortality and morbidity be lowered as much as possible during pregnancy and, in particular, in labor. Since the fetal heart rate was in the field of complex systems, non-linear methods are at the center of future research for good prediction of poor neonatal outcomes. However, the non-linear indexes must be understood in the light of the well-known linear measurements of fetal heart rate analysis commonly used in clinical practice. This thesis relates different parameters and supports a multiparameter approach in the interpretation of the fetal heart rate signal for acidemia detection and labor prediction.

References

1. Ayres-de-Campos, D., et al., *SisPorto 4.0 - computer analysis following the 2015 FIGO Guidelines for intrapartum fetal monitoring.* J Matern Fetal Neonatal Med, 2017. **30**(1): p. 62-67.

2. Marzbanrad, F., et al., *Transfer entropy analysis of maternal and fetal heart rate coupling.* Conf Proc IEEE Eng Med Biol Soc, 2015. **2015**: p. 7865-8.

3. Fulcher, B.D., et al., *Highly comparative fetal heart rate analysis.* Conf Proc IEEE Eng Med Biol Soc, 2012. **2012**: p. 3135-8.

4. Choi, W.Y. and J.K. Hoh, *Nonlinear analysis of fetal heart rate dynamics in fetuses compromised by asymptomatic partial placental abruption.* Placenta, 2015. **36**(12): p. 1474-9.

5. Spilka, J., et al., *Nearest-neighbor based wavelet entropy rate measures for intrapartum fetal heart rate variability.* Conf Proc IEEE Eng Med Biol Soc, 2014. **2014**: p. 2813-6.

6. Magenes, G., et al., *Multivariate analysis based on linear and non-linear FHR parameters for the identification of IUGR fetuses.* Conf Proc IEEE Eng Med Biol Soc, 2014. **2014**: p. 1868-71.

7. Tetschke, F., et al., *Assessment of fetal maturation age by heart rate variability measures using random forest methodology.* Comput Biol Med, 2016. **70**: p. 157-62.

8. Monteiro-Santos, J., et al., *Entropy and Compression Capture Different Complexity Features: The Case of Fetal Heart Rate.* Entropy, 2017. **19**(12).

9. Zarmehri, M.N., et al., *On the prediction of foetal acidaemia: A spectral analysis-based approach.* Computers in Biology and Medicine, 2019. **109**: p. 235-241.

10. Monteiro-Santos, J., et al., *Complexity of Cardiotocographic Signals as A Predictor of Labor.* Entropy (Basel, Switzerland), 2020. **22**(1): p. 104.

11. Elliott, C., et al., *Graded classification of fetal heart rate tracings: association with neonatal metabolic acidosis and neurologic morbidity.* Am J Obstet Gynecol, 2010. **202**(3): p. 258 e1-8.

12. Pincus, S.M. and R.R. Viscarello, *Approximate entropy: a regularity measure for fetal heart rate analysis.* Obstet Gynecol, 1992. **79**(2): p. 249-55.

13. Strachan, B.K., et al., *Computerised analysis of the fetal heart rate and relation to acidaemia at delivery.* BJOG, 2001. **108**(8): p. 848-52.

14. Chung, T.K., et al., *The prediction of fetal acidosis at birth by computerised analysis of intrapartum cardiotocography.* Br J Obstet Gynaecol, 1995. **102**(6): p. 454-60.

15. Pasarica, A., et al., *Dynamic of couplings between fetal heart rate and uterine contractions.* 2015 International Symposium on Signals, Circuits and Systems (Isscs), 2015.

16. Georgoulas, G., C. Stylios, and P. Groumpos, *Feature extraction and classification of Fetal Heart Rate using wavelet analysis and Support Vector Machines.* International Journal on Artificial Intelligence Tools, 2006. **15**(3): p. 411-432.

17. Goncalves, H., et al., *Frequency domain and entropy analysis of fetal heart rate: appealing tools for fetal surveillance and pharmacodynamic assessment of drugs.* Cardiovasc Hematol Disord Drug Targets, 2008. **8**(2): p. 91-8.

18. Henriques, T., et al., *Entropy and compression: two measures of complexity.* J Eval Clin Pract, 2013. **19**(6): p. 1101-6.

19. Annunziata, M.L., et al., *Computerized analysis of fetal heart rate variability signal during the stages of labor.* J Obstet Gynaecol Res, 2016. **42**(3): p. 258-65.

20. Spilka, J., et al., *Using nonlinear features for fetal heart rate classification.* Biomedical Signal Processing and Control, 2012. **7**(4): p. 350-357.

21. Costa, M.D., et al., *Complexity-loss in fetal heart rate dynamics during labor as a potential biomarker of acidemia.* Early Hum Dev, 2014. **90**(1): p. 67-71.

22. Barquero-Perez, O., et al., *Fetal Heart Rate Complexity Measures to Detect Hypoxia.* 2015 Computing in Cardiology Conference (Cinc), 2015. **42**: p. 133-136.

23. Granero-Belinchon, C., et al., *Information Theory to Probe Intrapartum Fetal Heart Rate Dynamics.* Entropy, 2017. **19**(12).

24. Santiago-Mozos, R., et al., *On the Early Detection of Perinatal Hypoxia with Information-Theory based Methods.* 2013 Computing in Cardiology Conference (Cinc), 2013. **40**: p. 425-428.

25. d'Aloja, E., et al., *Neonatal asphyxia and forensic medicine.* J Matern Fetal Neonatal Med, 2009. **22 Suppl 3**: p. 54-6.

26. Heintz, E., et al., *The long-term cost-effectiveness of fetal monitoring during labour: a comparison of cardiotocography complemented with ST analysis versus cardiotocography alone.* BJOG, 2008. **115**(13): p. 1676-87.

27. Fukushima, T., et al., *Limitations of Auto-Correlation in Fetal Heart-Rate Monitoring.* American Journal of Obstetrics and Gynecology, 1985. **153**(6): p. 685-692.

28. Goncalves, H., et al., *Internal versus external intrapartum foetal heart rate monitoring: the effect on linear and nonlinear parameters.* Physiol Meas, 2006. **27**(3): p. 307-19.

29. Chudacek, V.S., Jiri; Huptych, Michal; Georgoulas, George; Janku, Petr; Koucky, Michal; Stylios, Chrysostomos; Lhotska, Lenka, *Automatic Classification of Intrapartal Fetal Heart-Rate Recordings - Can It Compete with Experts?* 2010.

30. Dawes, G.S., M. Moulden, and C.W.G. Redman, *Criteria for the Design of Fetal Heart-Rate Analysis Systems.* International Journal of Bio-Medical Computing, 1990. **25**(4): p. 287-294.

31. Bernardes, J., et al., *The Porto system for automated cardiotocographic signal analysis.* J Perinat Med, 1991. **19**(1-2): p. 61-5.

32. Goncalves, H., et al., *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**(6): p. 665-76.

33. Ayres-de Campos, D., et al., *SisPorto 2.0: a program for automated analysis of cardiotocograms.* J Matern Fetal Med, 2000. **9**(5): p. 311-8.

34. Chourasia, V.S.T., Anil Kumar, *FETAL HEART RATE VARIABILITY ANALYSIS FROM PHONOCARDIOGRAPHIC RECORDINGS.* 2011.

35. Signorini, M.G., et al., *Advances in monitoring cardiovascular signals. Contribution of nonlinear signal processing.* Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 6568-71.

36. Signorini, M.G., A. Fanelli, and G. Magenes, *Monitoring fetal heart rate during pregnancy: contributions from advanced signal processing and wearable technology.* Comput Math Methods Med, 2014. **2014**: p. 707581.

37. Khandoker, A.H., et al., *Analysis of maternal-fetal heart rate coupling directions with partial directed coherence.* Biomedical Signal Processing and Control, 2016. **30**: p. 25-30.

38. Khandoker, A.H., et al., *Alterations in Maternal-Fetal Heart Rate Coupling Strength and Directions in Abnormal Fetuses.* Frontiers in physiology, 2019. **10**: p. 482-482.