Multivariate Analysis based on Linear and Non-Linear FHR Parameters for the Identification of IUGR Fetuses

G. Magenes, *IEEE Member*, R. Bellazzi, A. Fanelli and M.G. Signorini, *IEEE Member*

Abstract— **Fetal Heart Rate (FHR) monitoring represents a powerful tool for checking the arousal of pathological fetal conditions during pregnancy. This paper proposes a multivariate approach for the discrimination of Normal and Intra Uterine Growth Restricted (IUGR) fetuses based on a small set of parameters computed on the FHR signal. We collected FHR recordings in a population of 120 fetuses (60 normals and 60 IUGRs) at approximately the same gestational week through a standard CTG non-stress test. A set of 8 linear and non-linear indices were selected and computed on each recording, on the basis of their "stand-alone" discriminative properties, demonstrated in previous studies. By using the Orange® data mining suite we checked various multivariate discrimination models. The results show that a Logistic Regression performed on a limited set of only 4 parameters can reach 92.5% accuracy in the correct identification of fetuses, with 93% sensitivity and 91.5% specificity.**

I. Introduction

The history of Fetal Heart Rate (FHR) analysis has been strongly influenced by developments in signal processing methods. Since the introduction of Cardiotocography in the early '70ies, FHR monitoring represents an almost unique method to survey the fetal development in a noninvasive and quite simple way. Changes in fetal heart beats have been observed before any other sign of disease clearly appears, thus assuming a growing importance in the diagnostic process [1].

Starting from the first prototypes of computerized Cardiotocography, for a long time, the FHR signal was analyzed by classic time domain parameters in order to evaluate the variability in short and long time windows and to deduce possible pathological or risky conditions for the fetus from changes in the FHR variability [2]. Since the end of 80's, the introduction of non linear parameters in the analysis of biological signal dynamics reinforced even more this viewpoint by investigating the geometric and dynamic properties of heart rate time series. In fact, some aspects of the FHR variability that were never well understood, received a new attention thanks to the novel available tools [3].

G. Magenes is with the Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, 27100 Pavia, Italy, (corresponding author phone: +39 0382-985515, fax: 39 0382-985373; e-mail: giovanni.magenesr@ unipv.it).

R. Bellazzi is with the Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, 27100 Pavia, Italy, (e-mail: riccardo.bellazzzi@unipv.it).

Andrea Fanelli is with the Computational Physiology and Clinical Inference Group, MIT, Cambridge, MA, USA, (e-mail: fanelli@mit.edu)

M.G. Signorini is with the Dipartimento di Elettronica Informazione e Bioingegneria-DEIB, Politecnico di Milano, Italy, (e-mail: mariagabriella.signorini@polimi.it)

However, for many years, the research in this field has been focused to find a single indicator that could be the winner in discriminating healthy from pathological fetuses. Only recently, developments in the biosignal analysis research and application provided a novel view in the FHR analysis and in its possible diagnostic use. Instead of searching for single parameters, researchers started to understand that the problem was complex and required more than one index to be correctly faced. Experimental results demonstrated that many mechanisms were acting in the signal control in particular when disease conditions arise [4].

The multiparametric approach proposed in this manuscript is the result of a research work, lasting since the beginning of the new millennium, dedicated to extract diagnostic information from FHR signals collected during pregnancy. The goal is to find a limited set of linear and nonlinear parameters that, by combining different methods of parametrization of FHR signal, can generate a new classification model for the early discrimination of IUGR fetuses.

II. Methods

A. Data collection and preprocessing

FHR signals were collected on a population of 120 pregnant women at the Azienda Ospedaliera Universitaria Federico II, Napoli, Italy, through a Hewlett Packard CTG fetal monitor, connected with a PC computer. The set of recordings was composed of 60 healthy and 60 IUGR fetuses. Both groups were identified after delivery, on the basis of Apgar scores, weight and abdominal circumference at birth: IUGR fetuses were selected by weight below the 10th percentile for their gestational age, abdominal circumference below the 10th percentile and Apgar score <8. The CTG recordings were performed in a controlled clinical environment, with the pregnant woman lying on a bed. The details of the two populations are reported in Table I. The average length of the recordings was 2742 ± 595 sec for healthy subjects and 3412 ± 1023 sec for IUGR fetuses.

TABLE I – Summary of Normal and IUGR fetuses

Population	Healthy	IUGR
Number	60	60
Mother (ys)	32,34±5,64	29.68±6.21
Gest. age at CTG (ws)	34.78 ± 0.53	32,27±2,79
Gest. age at birth (ws)	39,74 ± 1,15	34,15 ± 2,99
Weight of the baby	3275 g± 518g	1479 g ± 608 g
Delivery mode	58% Spont. 42% Caesarean	14,8% Spont. 85,2% Caesarean

As described in [3] the HP fetal monitors use an autocorrelation technique to compare the demodulated Doppler signal of a heartbeat with the next one and a peak position interpolation algorithm, which allow an effective resolution better than 2 ms in the detection of the heart period. The HP monitor computes FHR values every 250 msec and, in the commercially available system, it provides the actual FHR in bpm every 2.5 sec as the average of 10 consecutive values (corresponding to an equivalent sampling frequency of 0.4 Hz). In our system we modified the software in order to read the FHR at 2 Hz (every 0.5 sec), which represents a reasonable compromise to achieve an enough large bandwidth (Nyquist Frequency 1 Hz) and an acceptable accuracy of the FHR signal.

B. Linear FHR Parameters

Classical linear indices are usually computed on the time course of the FHR signal by excluding accelerations and decelerations as proposed by Arduini et al. [5]. Among these classical indices we calculated the Short Term Variability (STV), the Long Term Irregularity (LTI), and the covariance of FHR signal (RCO).

STV quantifies FHR variability over a very short time scale. We refer to definitions provided by Dalton et al. [6] (even if we used a scale factor of 12) and by Arduini et al. [5]. By considering one minute of interbeat sequence, we define STV as:

$$STV = mean\left[\left|T_{24}(i+1) - T_{24}(i)\right|\right]_i = \frac{\sum_{i=1}^{23}\left|T_{24}(i+1) - T_{24}(i)\right|}{23}$$

where $T_{24}(i)$ is the value of the FHR signal in milliseconds, taken each 2.5 sec. STV was computed for each minute and the averaged on the whole recording.

LTI was proposed by De Haan et al. [7]. It is computed on a three minute segment of interbeat sequence $T(i)$ in milliseconds, by excluding accelerations and decelerations. LTI is defined as the interquartile range [¼; ¾] of the distribution $m(j)$ with j belonging to a three minute window, $m(j)=\sqrt{T(j)^2+T(j+1)^2}$; where $T(j)$ is the FHR signal in milliseconds. LTI values were then averaged on the whole recording.

C. Non -Linear FHR Parameters

Various techniques exist aimed at quantifying the degree of similarity and/or complexity in time series. Among them we considered Approximate Entropy (ApEn), Sample Entropy (SampEn), Lempel Ziv Complexity (LZC).

ApEn [8] and SampEn [9] both quantify regularity and complexity of a time series. They evaluate the signal regularity, within a tolerance r, by assessing the frequency of patterns similar to a pattern of window length m (m=1, 2, r: 0.1 - 0.25 std of the input data).

SampEn has been introduced and largely employed in biomedical signal processing over time, as it improves the estimation performed by ApEn (i.e. removes the bias introduced by self-counts). In our experimental data ApEn and SampEn were estimated on the FHR time series, on non-overlapping windows of three minutes, by using the same parameter set: m=1 and r=0.1, m=2 and r=0.15 and 0.2.

Lempel Ziv Complexity (LZC) was introduced in the field of Information Theory to measure the number of different sub strings and the rate of their recurrence [10]. Namely, LZC reflects the gradual increase of new patterns along a given sequence. The measure of complexity introduced by Lempel and Ziv is defined as the minimum quantity of information needed to define a binary string. In order to estimate the LZC in a time series, it is necessary to transform the signal (the FHR in our case) into symbolic sequences. We transformed FHR signals through a binary and a ternary coding procedures. In the binary coding, given a FHR series $\{x_n\}$, we built the sequence y_n by assigning 1 to a signal increase ($x_{n+1}>x_n$) and 0 to a decrease ($x_{n+1}\leq x_n$). In case of ternary alphabet, 1 denotes the signal increase ($x_{n+1}>x_n$), 0 the decrease ($x_{n+1}<x_n$) and 2 the signal invariance ($x_{n+1}=x_n$). To avoid the possible dependence of the encoded string on quantization procedure adopted to record the signal, a p factor is introduced representing the minimum quantization level for a symbol change in the coded string (e.g. $y_n = 1$ if $x_{n+1}>x_n+x_n.p$). We considered the encoding parameter p = 0, 0.005, 0.01, 0.02. The LZC index was computed on 360 point-long FHR windows (3 min).

In addition to complexity or regularity indices, we recently introduced two novel parameters computed on the Phase Rectified Signal Average (PRSA) curves, namely the Acceleration Phase Rectified Slope (APRS) and the Deceleration Phase Rectified Slope (DPRS) [11]. Phase Rectified Signal Average (PRSA) was proposed by Bauer et al. [12] and allows the detection and quantification of quasi-periodic oscillations in non-stationary signals affected by noise and artifacts, by synchronizing the phase of all periodic components to "anchor points". This method is useful to enhance episodes of increasing and/or decreasing FHR, which are functionally related to fetal condition. The procedure to build up the PRSA curve is described in [12]. An interesting feature of a PRSA curve is that a 30-40 minutes FHR signal can be condensed in a single waveform, showing the average dynamic pattern of the recording under analysis. For each FHR recording we built two PRSA curves, by taking 200 sec windows (400 samples) from the FHR signal, which were selected if the right average of the window was higher/lower than the left average. Then, the windows were synchronized in their anchor points (the middle point of the curve) and averaged. In order to summarize the behavior of each PRSA curve with one figure of merit, we proposed the APRS and DPRS, namely the slope of increasing/decreasing PRSA curves in the anchor point.

TABLE II – SUMMARY OF THE RESULTS FOR THE CHOSEN PARAMETERS

Parameter	Healthy (mean ± std)	IUGR (mean ± std)	t-test	p-value
Time Domain				
Rcov(0) (ms²)	349 ±115	175 ±72	***	2.10 e-16
STV (ms)	6.7 ± 2.24	4.29 ± 1.62	***	1,22 e-09
LTI (ms)	21,46 ± 6.53	17.17 ± 5.37	***	1.5 e-11
Regul./Compl.				
ApEn(1,0.1)	1.33 ± 0.13	1.21 ± 0.11	**	5.14 e-7
Lempel Ziv (2,0)	1.00±0.08	0.94±0.09	*	0.00078
SampEn(1,0.1)	1.3±0.19	1.13±0.15	**	2.08 e-7
PRSA				
APRS	0.17± 0.041	0.12 ±0.042	***	7.76 e-12
DPRS	-0.18± 0.046	-0.12 ±0.042	***	1.08 e-13

D. Multivariate Analysis

The data have been analyzed with the Orange software tool, a Python data-mining suite implemented by the University of Ljubljana [13]. A set of multivariate methods, including Logistic Regression (LR) with stepwise variable selection, Naïve Bayes (NB), Support Vector Machines with Linear and Gaussian Kernel, Classification trees, have been tested within a ten-fold cross-validation scheme. The performance of the method has been assessed by computing Accuracy, Sensitivity, Specificity, Area under the ROC curve (AUC), F-measure, Brier Score and Matthews Correlation Coefficient. Finally, a simple scoring system has been extracted from the entire data set by applying stepwise variable selection to a LR model.

III. Results

The accuracies of all the tested methods, shown in Table III, were higher or equal than 90%, with the exception of the Classification Tree.

In general, the data are well separated by a linear decision boundary, as shown by all performance indices. In particular, both SVM with a linear kernel and LR show the same performance in terms of accuracy, sensitivity and specificity, but LR obtains a better AUC. For these reasons, the LR classifier was selected, as it provides: i) an easy to be interpreted classification model, based on an additive scoring system of the problem variables; ii) statistical tools to select the most important variables.

The ROC curve of the LR model is shown in Figure 1.

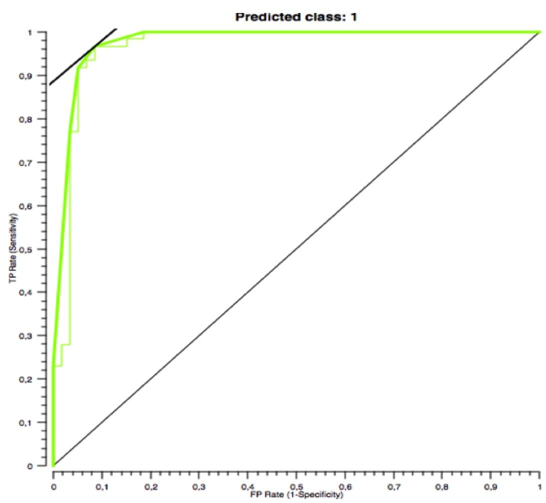

Fig. 1 - The empirical ROC of the logistic regression model .

TABLE III Performance of the Multivariate Classification Method

	Method	CA	Sens	Spec	AUC	F1	Brier	MCC
1	Classification Tree	0.8417	0.8361	0.8475	0.8522	0.8430	0.2762	0.6834
2	Naive Bayes	0.9000	0.8852	0.9153	0.9806	0.9000	0.1576	0.8005
3	Logistic regression	0.9250	0.9344	0.9153	0.9806	0.9268	0.1187	0.8500
4	SVM - RBF	0.9083	0.9508	0.8644	0.9722	0.9134	0.1575	0.8192
5	SVM - Linear	0.9250	0.9344	0.9153	0.9778	0.9268	0.1166	0.8500

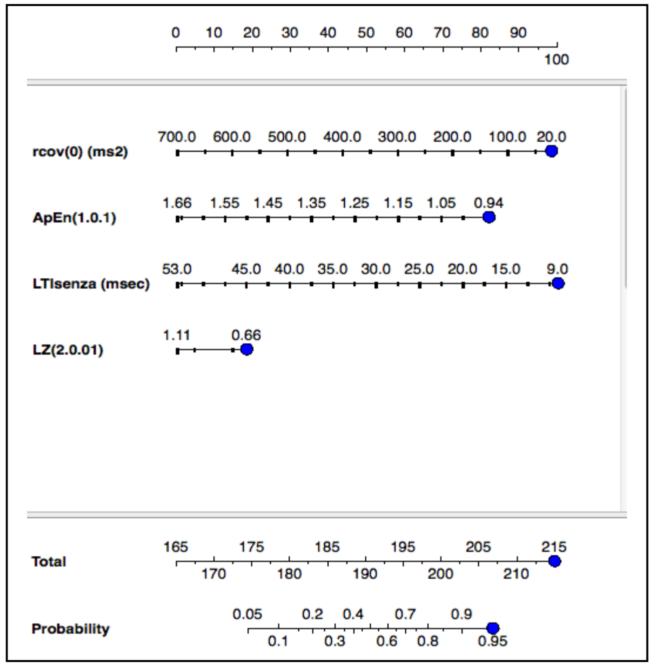

Fig. 2 - The nomogram derived from the LR model. The value of each variable is associated to a score. A total score is computed as the sum of the scores of the variables. The total score is then mapped to the probability of being a IUGR fetus. In the example, a subject with rcov=20, Apen=0.94, LTI=9 and LZ=0.66 has a total score of 215 and a probability of being a IUGR fetus equal to 0.95.

When compared with the LR built with the best univariate predictor, covariance, the multiparametric strategy allows significantly improving the 10-fold prediction performance in terms of accuracy (from 84% to 92%), with a 10% improvement of sensitivity.

After running the stepwise selection algorithm with the LR model, four variables have been retained: Apen, LTI, LZC and RCO. These variables have negative odds ratios. Therefore, all variables values are inversely related to the probability of being an IUGR fetus. Thanks to the model, it is possible to derive a simple scoring system that can be represented with a nomogram, as shown in Figure 2.

IV. Discussion and Conclusion

The problem of identifying IUGR fetuses during pregnancy has been faced through various approaches in the literature. In the recent years our group proposed different indices based on LZC and multiscale entropy [14,15] and on DPRS and APRS [11]. Other researchers introduced the idea of Power Rectified Signal Average and used different indices (average acceleration capacity - AAC) [16, 17]. Most of these parameters are highly significant in discriminating healthy fetuses from IUGRs, but all considered a single feature extracted from the FHR signal and tried to demonstrate that the proposed index was better than the others.

The truth is that a single index cannot summarize by itself the features of all pathophysiological processes driving the development of a fetus towards the IUGR condition. In previous works we examined couples of indices together in order to check if clusters exist able to separate healthy from IUGR fetuses in a two-dimensional domain [4, 15], obtaining in both cases quite good results.

As a matter of fact, many controlling mechanisms can affect the heart rate variability and they may act linearly and/or non linearly on the FHR in pathological situations. Only a multivariate approach, considering both linear and non linear parameters, can really improve the discrimination of healthy and pathological fetuses.

In this paper we used multivariate analysis of FHR parameters with an almost complete, although not exhaustive, approach. We selected a set of eight parameters (STV, LTI, covariance, ApEn, SampEn, LZC, APRS and DPRS), three of them belonging to classical linear indices and five related to non linear properties of the FHR signal, all showing discriminative ability for IUGR fetuses when adopted as "stand-alone" parameters [4].

By means of data mining techniques available in Orange, we exploited some of the most popular classification methods, able to deal with both continuous and discrete variables. We tested methods designed to learn linear and non-linear decision boundaries. In case of similar performance, we selected methods that provide transparent classification rules.

In particular, we were able to show that a simple Logistic Regression model, using only four of our parameters (covariance, LTI, ApEn and LZ Complexity) can reach very high performance in the discrimination task, much higher than any single parameter by itself. In the four-dimensional space of the parameters healthy and IUGR populations are separated by a linear decision boundary, which does not happen if we consider the parameters in a lower dimensional space. Moreover, the reduced set of four parameters consists of two linear and two non linear indices, implicitly showing that different control mechanisms play important roles in the development of a pathological condition.

It is important to mention that the four parameters have been chosen on the basis of the well-known stepwise procedure, which is designed to select variables combining both backward and forward elimination. This process tends to select parameters that are not correlated and that provide good classification performance when they are jointly used. This of course does not mean that other parameters, dropped from the model only because they are correlated with the selected ones, have no clinical importance.

In conclusion, our multiparametric approach should be extensively tested in the clinical practice. Its main advantage resides in the fact that it can produce an overall score extracted from the four selected parameters and the physician can immediately interpret this score.

Acknowledgment

The authors are grateful to Prof. A. Di Lieto and Dr. M. Campanile, Obstetrics and Gynaecology Department of the Università Federico II, Naples, Italy for the collection of CTG tracings on healthy and IUGR fetuses.

References

[1] H. P. van Geijn, "Developments in CTG analysis," Baillieres Clin. Obstet. Gynaecol., vol. 10, no. 2, pp. 185–209, 1996.

[2] G. S. Dawes, M. Moulden, and C.W. Redman, "The advantages of computerized fetal heart rate analysis," J. Perinat. Med., vol. 19, no. 1–2, pp. 39–45, 1991.

[3] 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 Biom Eng, 50(3), 365-375, 2003..

[4] M.G. Signorini, A. Fanelli, G. Magenes "Monitoring Fetal Heart Rate during Pregnancy: Contributions from Advanced Signal Processing and Wearable Technology" Computational and Mathematical Methods in Medicine. 03/2014; Vol. 2014:10 pages, DOI:10.1155/2014/707581.

[5] D. Arduini, G. Rizzo, G. Piana, A. Bonalumi, P. Brambilla, and C. Romanini, "Computerized analysis of fetal heart rate: I. description of the system (2CTG)," Matern. Fetal. Invest. 3, 159, 1993.

[6] K. J. Dalton, G. S. Dawes, and J. E. Patrick, "Diurnal, respiratory and other rhythms of fetal heart rate in lambs," Am. J. Obstet. Gynecol. 127, 414, 1977.

[7] J. De Haan, J. H. Vam Bemmel, B. Versteeg, A. F. L. Veth, L. A. M. Stolte, J. Janssens, and T. K. A. B.Eskes, "Quantitative evaluation of fetal heart rate. i. processing methods" Eur. J. Obstet. Gynaecol. 3, 95, 1971.

[8] S. M. Pincus, "Approximate entropy (ApEn) as complexity measure", Chaos, 5(1), 110-117, 1995.

[9] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy", Am J Physiol Heart Circ Physiol, 278, 2039-2049, 2000.

[10] A. Lempel and J. Ziv, "On the complexity of finite sequences", IEEE Transactions on Information Theory , 22, 75-81, 1976.

[11] A. Fanelli, G. Magenes, M. Campanile, M.G. Signorini, "Quantitative assessment of fetal well-being through CTG recordings: a new parameter based on Phase Rectified Signal Average", IEEE Journal of Biomedical and Health Informatics, Sept 2013, Vol 17 (5), 959–966, 2013.

[12] G.S. A. Bauer, A. Bunde, P. Barthel, R. Schneider, M. Malik, G. Schmidt, "Phase rectified signal averaging detects quasi-periodicities in non-stationary data", J Phys A, vol. 364, pp. 423–434, 2006.

[13] L. Demšar, T. Curk, A. Erjavec. Orange: Data Mining Toolbox in Python; Journal of Machine Learning Research 14(Aug):2349−2353, 2013.

[14] M. Ferrario, M.G. Signorini, G. Magenes, "Comparison between fetal heart rate standard parameters and complexity indexes for the identification of severe intrauterine growth restriction.", Methods of Information in Medicine , 46 (2), 186-190, 2007.

[15] M. Ferrario, M.G. Signorini, G. Magenes. "Complexity analysis of the fetal heart rate variability: early identification of severe intrauterine growth-restricted fetuses". Med Biol Eng Comput. 47(9), 911-919, 2009

[16] E. A. Huhn, S. Lobmaier, T. Fischer, R. Schneider, A. Bauer, K. Schneider and G. Schmidt, "New computerized fetal heart rate analysis for surveillance of intrauterine growth restriction", Prenat Diagn, nn. 31, vol. 5, pp. 509-514, 2011.

[17] Lobmaier SM, Huhn EA, Pildner von Steinburg S, Müller A, Schuster T, Ortiz JU, Schmidt G, Schneider KT. Phase-rectified signal averaging as a new method for surveillance of growth restricted fetuses. J Matern Fetal Neonatal Med. 2012 Dec;25(12):2523-8..