fetal monitoring, fetal heart rate,
signal classification, ROC analysis

Robert CZABAŃSKI[1], Michał JEŻEWSKI[1], Dawid ROJ[2],
Zbigniew SZASZKOWSKI[3], Tomasz KUPKA[2], Janusz WRÓBEL[2]

EVALUATION OF PREDICTIVE CAPABILITIES OF QUANTITATIVE CARDIOTOCOGRAPHIC SIGNAL FEATURES

Cardiotocography (CTG) is the main method of assessment of the fetal state during pregnancy and labour used in clinical practice. It is based on quantitative analysis of fetal heart rate, fetal movements and uterine contractions signals. The evaluation of the CTG signals can be made using criteria recommended by International Federation of Obstetrics and Gynecology. Nevertheless, the diagnosis verification is possible only after the delivery on the basis of newborn assessment. In the proposed work we evaluated the capacity of quantitative analysis of CTG traces in predicting fetal outcome. The relationship between CTG signal features and attributes of fetal outcome was assessed on the basis of ROC curves analysis. The obtained results indicate the adequate predictive capabilities of the selected CTG features especially for fetal outcome assessed with Apgar score and suggest the necessity of applying the criteria for the CTG traces evaluation that are related to the gestational age. Our study also shows the value of the CTG monitoring as a screening procedure providing appropriate confirmation of fetal wellbeing.

1. INTRODUCTION

Cardiotocographic (CTG) monitoring is a primary biophysical method of assessing the condition of a fetus during pregnancy and labour. It involves simultaneous recording and analysis of three signals: fetal heart rate, contractile activity of uterine muscle and fetal movements. The fetal heart rate (FHR) is commonly regarded as of primary diagnostic significance. Nevertheless, the visual analysis of signal patterns describing the FHR variability is extremely difficult due to the complexity of waveforms shape. At the same time, the FHR signal contains a number of diagnostically relevant information that are hidden for the naked eye and can only be quantitatively described with computer analysis. Modern systems for computer-aided cardiotocography offer the automated quantitative description of the CTG records, but the development of methods ensuring effective support for the qualitative assessment are is the aim of many studies [3,4,7]. Numerous attempts were made to formalize the clinical guidelines for the interpretation of CTG recording on the basis of its quantitative analysis. Nevertheless, the criteria provided by the International Federation of Obstetrics and Gynecology [5] (FIGO fr. Fédération Internationale de Gynécologie et d'Obstétrique) defined the standards used in the perinatal medicine.

During the CTG monitoring we evaluate the actual (at the time of recording session) fetal state. Unfortunately it is not possible to verify the evaluation result at the time of the CTG monitoring using any other reference technique. The information is to be revealed after the delivery only. Nevertheless, the prediction of fetal outcome during pregnancy is possible, because in perinatology it is assumed that the fetal state cannot change rapidly. Therefore, the newborn state just after delivery can be retrospectively assigned to the condition of the fetus at the time of the CTG recording. The aim of this study is to investigate the possibility of fetal state prediction based on cardiotocographic signal features evaluation in accordance with FIGO guidelines.

2. FIGO CRITERIA

FIGO criteria for fetal assessment are based on the most important quantitative parameters of FHR signal including: changes of the basal fetal heart rate (baseline), changes of the fetal heart rate in certain

[1] Silesian University of Technology, Institute of Electronics, ul. Akademicka 16, 44-100 Gliwice, Poland,
[2] Institute of Medical Technology and Equipment, Biomedical Informatics Department, ul. Roosevelta 118, 41-800 Zabrze, Poland,
[3] Silesian University of Technology, Institute of Mathematics, ul. Kaszubska 23, 44-100 Gliwice, Poland.

direction, e.g. transitory increases (accelerations, ACC) and decrease (decelerations) of the FHR as well as short-lasting changes of the fetal heart rate also called instantaneous variability. The antenatal FHR patterns are divided into three classes indicating the fetal state as “normal”, “suspicious” or “pathological”.

According to FIGO, the baseline FHR is defined as “the mean level of the FHR when this is stable, accelerations and decelerations being absent. It is determined in the time interval from 5 to 10 min and expressed in beats per minute." [5]. Accelerations and decelerations are defined as temporary deviations of the FHR in relation to the baseline, with a certain range of amplitude and duration. The acceleration is recognized if the increase in FHR above the baseline is of 15 bpm or more and lasting 15 s or more. Decelerations are defined as decreases of the FHR under the baseline with the amplitude exceeding 15 bpm and lasting minimum of 10 s. There are three types of deceleration distinguished in the MONAKO system [6]: D_A (> 15 bpm, > 10 s), D_B (> 10 bpm, > 25 s) and D_C (> 15 bpm, > 10 s). For the patterns of type C an additional criterion is to associate the deceleration with the contraction of the uterus. We extended the FIGO criteria for all types of recognized decelerations.

There are two types of instantaneous FHR variability being considered: short-term (STV), defining changes of intervals between two consecutive heart beats, and long-term variability (LTV) with periodical changes of short-term variability concerning both direction and magnitude, resulting in FHR oscillations. In our approach, the parameter STV is determined within a minute as an average value of the absolute difference of two adjacent FHR samples [9]. LTV parameter is determined for a period of one minute as the difference between maximum and minimum value of the averaged FHR samples. There are four types of oscillations distinguished in the MONAKO system [6]:

- type III - saltatory oscillation, amplitude exceeding 25 bpm,
- type II - undulatory oscillation, amplitude from 10 to 25 bpm,
- type I - a narrow undulatory oscillation, amplitude of 5 to 10 bpm,
- type 0 - silent oscillation, amplitude of less than 5 bpm.

We extended the FIGO recommendations for the parameters of instantaneous variability as well. As a criterion for oscillations we used their percentage in a whole trace. The values ranges indicating the classes of the fetal state are shown in Table 1.

Table 1. The classification of antepartum FHR signals according to FIGO guidelines.

Quantitative parameter	Normal	Suspicious	Pathological
Baseline [bpm]	[110, 150]	[100, 110) or (150, 170]	[0,100) or > 170
STV [ms]	[6, 14]	> 14	[0, 6)
LTV [ms]	>32	[24, 32]	<24
Accelerations [number per hour]	> 12	(1.5, 12]	[0, 1.5]
Decelerations [number per hour]	$D_A \in [0, 1.5)$ and $D_B = 0$ and $D_C = 0$	$D_A \geq 1.5$ or $D_B \in [0, 1.5)$ or $D_C \in [0, 1.5)$	$D_B \geq 1.5$ or $D_C \geq 1.5$
Oscillations [%]	$O_0 = 0$ and $O_I \in [0, 40)$ and $O_{III} = 0$	$O_0 \in [0, 40)$ and $O_I \geq 40$	$O_0 \geq 40$

Since there is no diagnostic method providing the reference information about the true fetal state at the time of CTG monitoring, the diagnosis verification is possible only after the delivery on the basis of fetal outcome assessment. The fetal outcome is evaluated by clinicians usually with a help of four crucial attributes [8]: percentile of birth weight, Apgar score at the first, fifth and tenth minute of life, umbilical artery (pH) and umbilical artery base excess (BE) measurements. Similarly to FIGO criteria three ranges of values are associated with the condition of the newborn described as normal, suspicious and pathological. The classification of attributes characterizing the fetal state are presented in Table 2.

Table 2. The established ranges of attributes describing the fetal state.

Fetal outcome index	Normal	Suspicious	Pathological
Percentile of birth weight	≥ 10	[5,10)	< 5
Apgar score	≥ 7	5 or 6	< 5
Umbilical artery pH	≥ 7.2	[7.1, 7.2)	< 7.1
Umbilical artery base excess	> 12	> 10	> 0

3. METHODOLOGY

The research material used in our study contains a set of quantitative parameters of CTG signals recorded using the computerized fetal surveillance system MONAKO (ITAM Zabrze) combined with medical data referring to particular patients and their newborns. After excluding records characterized by acquisition time less than 20 minutes or signal loss more than 20%, 2125 traces from 328 patients (newborns) were qualified as the final database.

One pregnant woman can be monitored many times before delivery, especially in case of high-risk pregnancy. With the progress of pregnancy the features characterizing the CTG signals are significantly changing. Therefore, we decided to use three ways of CTG traces selection. In the first, we used all available records (multiple CTG records), with the mean gestational age 36 weeks, median 36, minimum and maximum 24 and 44 weeks respectively. In the second, the earliest patient's CTG traces (single CTG records) were used (but before 37th week of gestation). For this dataset, the mean and median gestational age at the time of the CTG recording was 33 weeks and the minimum 24. In the third approach we used the last CTG traces recorded a few hours before the labour started. The average and median gestational age was in this case 39 weeks (the maximum 44, the minimum 31).

The fetal state was determined based on the single fetal outcome attribute value or with their combination in the form of a logical sum (LS), where the fetal outcome was defined as abnormal if the value of at least one attribute was outside the physiological range. The available newborn data forms not always contained all values of fetal outcome attributes. Hence, the number of the CTG records analyzed varied depending on the choice of the attribute considered. We evaluated the predictive capabilities of quantitative FHR signal parameters with a help of receiver operating characteristic (ROC) analysis. Therefore, at first we included the "suspicious" CTG records into the class indicating the normal fetal state and next into the class representing the abnormal fetal state. Table 3 shows structures of data used in our experiments. The least number of records was obtained for evaluation of the fetal state based on logical sum of the assessments of individual fetal outcome attributes. It is due to the requirement of complete newborn data forms.

Table 3. The structures of data used in our experiments.

Fetal outcome index	Number of CTG records		
	Multiple	Single	Last
Birth weight	1944 : 260 : 0 *)	85 : 48 : 0	123 : 10 : 0
1-min Apgar score	2115 : 465 : 162	120 : 31 : 15	156 : 51 : 24
5-min Apgar score	2124 : 175 : 74	121 : 12 : 7	157 : 16 : 4
10-min Apgar score	2124 : 155 : 27	121 : 11 : 5	157 : 13 : 2
Umbilical artery pH	1770 : 204 : 63	121 : 15 : 5	142 : 29 : 5
Umbilical artery base excess	1739 : 167 : 93	89 : 12 : 5	133 : 17 : 8
Logical sum	1561 : 412 : 111	56 : 14 : 2	101 : 28 : 9

*) total number of CTG records in dataset : number of pathological cases if suspicious cases were assigned to the class pathological : number of pathological cases if suspicious cases were assigned to the class normal

To assess the capability of predicting the fetal outcome using quantitative analysis of CTG recording in accordance with the FIGO criteria we used ROC curves. They allow us to analyze the relation between the basic measures of the prediction performance (sensitivity SE and specificity SP), as the discrimination thresholds of the FHR parameters variation. The area under the ROC (AUC) is equal to

the probability that the rank of a randomly chosen instance for abnormal fetal outcome *X* is higher than a randomly chosen normal one *Y* [2]:

$$\mathrm{AUC} = \mathrm{Prob}(X > Y). \tag{1}$$

Hence, we can evaluate the diagnostic significance of the CTG quantitative parameters using AUC values. In order to make the assessment results more convenient, we applied the scale of the predictive capabilities presented in Table 4.

Table 4. Classification of the area under the ROC curve.

Prediction capabilities	Fail	Poor	Fair	Good	Excellent
AUC	[0.5, 0.6)	[0.6, 0.7)	[0.7, 0.8)	[0.8, 0.9)	[0.9, 1.0]

Areas under the ROC curves were calculated with the standard deviation (SD) and 95% confidence interval (CI). Additionally, we evaluated the best cut-off value (BCV), providing the highest values of prediction performance measures.

4. RESULTS

The results of our investigations are shown in Table 5, and they concern the database of multiple CTG records, assuming that the records with suspicious values of fetal outcome attributes are assigned to the "normal" class. Because there was no case of abnormal birth weight the analysis of the low birth weight risk was not possible. The obtained results show fair predictive ability of the fetal outcome with instantaneous FHR variability and the number of acceleration patterns. The LTV parameter is closely related to the oscillation percentage of each type and hence, the same conclusion can be extended in relation to the parameters O_0, O_I and O_{III}. The best predictive capabilities (AUC = 0.73) were obtained for saltatory oscillation O_{III}. The 5-min Apgar score and pH measurement are the fetal outcome attributes for which we got the highest prediction quality. However, there is no relationship between quantitative parameters of FHR and BE or LS.

Table 5. Area under the ROC curve for data containing multiple CTG traces.

Quantitative parameter	1-min Apgar	5-min Apgar	10-min Apgar	pH	BE	LS
FHR	0.52	0.52	0.55	0.54	0.57	0.55
STV	0.62	0.70	0.66	0.69	0.61	0.55
LTV	0.66	0.71	0.65	0.72	0.63	0.57
ACC	0.60	0.65	0.70	0.64	0.61	0.57
D_A	0.49	0.56	0.57	0.55	0.50	0.51
D_B	0.51	0.51	0.60	0.53	0.50	0.50
D_c	0.52	0.52	0.56	0.50	0.51	0.51
O_0	0.63	0.64	0.62	0.67	0.61	0.53
O_I	0.64	0.70	0.62	0.68	0.64	0.60
O_{III}	0.64	0.73	0.67	0.72	0.66	0.59

The detailed results of ROC analysis for parameters with the highest AUC are presented in Table 6. We obtained the same values of STV and ACC leading to the best distinction between normal and abnormal state of the fetus for both attributes of the fetal outcome. Values STV<5.249, ACC<5.300 and LTV<40.39 indicate the pathological fetal outcome when considering the Apgar score or pH measurement. These results are consistent with criteria according to FIGO guidelines however, in case of LTV, they are much rigorous when classifying fetal wellbeing.

When including the suspicious cases in the group of abnormal fetal state, we noticed the significant decrease of predictive capabilities for all quantitative parameters of FHR signal considered. The highest

value of AUC equal to 0.60 was obtained only for 1-min Apgar score with parameters LTV and ACC. When evaluating the risk of the low birth weight we got the highest value of AUC = 0.54 for parameters STV and O_I. These results indicate very poor relationship between the birth weight and the quantitative parameters of FHR signal when considering multiple CTG traces. The combination of suspicious and pathological cases resulted in reducing the differences between values of FHR signal parameters indicating the wellbeing and the health risk of the fetus respectively. These differences are so small that the possibility of separation of these classes is reduced, resulting in the observed AUV decrease.

Table 6. The detailed results of ROC analysis for selected FHR parameters.

	5-min Apgar					
	Multiple CTG traces		Single CTG traces		Day of delivery CTG traces	
	AUC	BCV	AUC	BCV	AUC	BCV
STV	$0.689 \pm 0.029^{0.746}_{0.632}$ *)	5.249	$0.647 \pm 0.074^{0.792}_{0.501}$	5.900	$0.626 \pm 0.180^{0.979}_{0.272}$	8.716
LTV	$0.706 \pm 0.029^{0.762}_{0.649}$	40.39	$0.731 \pm 0.059^{0.847}_{0.614}$	34.58	$0.716 \pm 0.054^{0.822}_{0.611}$	45.71
ACC	$0.649 \pm 0.027^{0.701}_{0.596}$	5.300	$0.682 \pm 0.066^{0.812}_{0.552}$	4.000	$0.542 \pm 0.043^{0.626}_{0.458}$	3.600
	pH					
STV	$0.685 \pm 0.034^{0.752}_{0.617}$	5.249	$0.715 \pm 0.160^{1.000}_{0.400}$	4.149	$0.593 \pm 0.186^{0.958}_{0.228}$	4.720
LTV	$0.722 \pm 0.032^{0.786}_{0.659}$	41.24	$0.751 \pm 0.104^{0.955}_{0.546}$	29.71	$0.632 \pm 0.152^{0.931}_{0.333}$	36.71
ACC	$0.638 \pm 0.031^{0.700}_{0.577}$	5.300	$0.733 \pm 0.085^{0.899}_{0.566}$	7.700	$0.713 \pm 0.083^{0.876}_{0.550}$	1.800

*) $\mathrm{AUROC} \pm \mathrm{SD}^{\max 95\%\,\mathrm{CI}}_{\min 95\%\,\mathrm{CI}}$

In the next experiment we used the database consisting of a single CTG record for one patient. The selection of the earliest records caused the significant increase of AUC for most quantitative parameters of FHR and all attributes of fetal outcome. The lowest values of AUC (with the mean, calculated for all attributes of fetal outcome, equal to 0.55) were obtained for deceleration rates. Nevertheless, for the FHR baseline the mean AUC was 0.64, and for BE and LS 0.62 and 0.61 respectively. Similarly to the previous experiment, the highest predictive capabilities were obtained for STV, LTV and ACC together with Apgar scores and pH measurement. For ACC and 10-min Apgar we got the AUC = 0.79. However, taking into account the mean values of AUC, the best prediction quality of these parameters was noticed for Apgar score at the fifth minute of life. The detailed results are presented in Table 6. The differences between the best discriminative values for both attributes of the fetal outcome can be noticed. Nevertheless, the obtained results are consistent with values ranges defined with FIGO criteria.

The combination of suspicious and pathological cases resulted in the reduction of predictive capabilities of all FHR signal parameters. The mean values of AUC decreased by 0.05. However, the results left higher in comparison to data containing multiple CTG records. We noticed also the increase of the low birth weight risk prediction quality (Table 7).

Table 7. The values of AUC for low birth weight prediction for data containing single CTG traces (suspicious cases assigned to the pathological class).

FHR	STV	LTV	ACC	D_A	D_B	D_c	O_0	O_I	O_{III}
0.52	0.60	0.63	0.63	0.55	0.55	0.54	0.60	0.65	0.60

In the last experiment we analyzed data consisting the CTG traces recorded on the day of delivery. Similarly to the previous example we obtained better results in comparison to the analysis of multiple CTG records. In contrast to the results of previous experiments we noticed the strongest relationship between deceleration patterns and 10-min Apgar score. We also obtained the highest mean AUC value of 0.77, and the maximum of 0.86 for decelerations type A (Table 8). Hence, we can conclude that deceleration patterns should be regarded as crucial signs of fetal wellbeing on the delivery day.

Table 8. The values of AUC for 10-min Apgar prediction with data containing the last CTG traces.

FHR	STV	LTV	ACC	D_A	D_B	D_c	O_0	O_I	O_{III}
0.60	0.63	0.72	0.54	0.86	0.78	0.68	0.66	0.58	0.64

For CTG traces recorded in the day of delivery we also noticed the decrease of predictive capabilities of STV, LTV and ACC. The detailed results are presented in Table 6. Despite STV parameter and 10-min Apgar, the best cut-off values remain consistent with the FIGO criteria defining the fetal wellbeing. However, the significant changes of their values can be noticed (especially for the STV and ACC). The obtained results confirm the necessity of applying different criteria for evaluating the CTG traces recorded during the childbirth. Analogously to the previous examples, the combination of records classified as suspicious with records indicating the risk of fetal health resulted in reduction of AUC. The mean values of AUC decreased on average by 12%.

A similar analysis of the capacity of quantitative analysis of CTG traces, that were recorded in the day of labour in predicting fetal outcome was presented in [1]. The results of our study confirm the rather moderate predictive capabilities of FHR signal parameters only when considering umbilical artery pH. We did not get, however, such a high discriminative capacity to predict Apgar score. The discrepancies are the result of differences in the criteria for selecting data being evaluated as the study presented in [1] involved only women scheduled for elective caesarean section, whose CTG recording was performed within 4 hours of delivery.

For all datasets considered, we noticed the relation between attributes of the fetal outcome and quantitative parameters of FHR signal with the highest values of AUC that is illustrated in Figure 1.

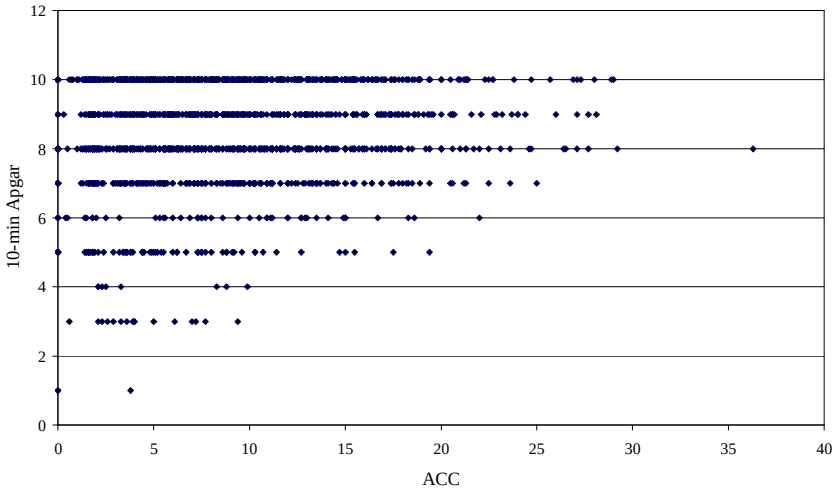

Fig. 1. Scatter-graphs of 10-min Apgar score relating to acceleration rate for the data with multiple CTG traces.

High Apgar scores and pH measurements indicating the wellbeing of the newborn are related to values of FHR signal parameters which, according to FIGO criteria, represent both normal and abnormal fetal state. However values of fetal outcome attributes outside the range indicating the newborn wellbeing are always associated with the CTG signal features indicating the fetal health risk. This relation proves the value of the CTG monitoring as a good screening procedure due to very high efficiency in confirmation of fetal wellbeing.

5. CONCLUSIONS

In the presented work, a number of experiments concerning the evaluation of the capacity of quantitative analysis of cardiotocographic traces in predicting fetal outcome were done. The relations between different attributes of newborn outcome and parameters of FHR signal distinguished by FIGO criteria were investigated. The predictive capabilities were evaluated on the basis of ROC curves analysis. The values of the areas under the ROC curves shows the relationship between FHR signals features and fetal outcome especially when assessed with Apgar score. The highest predictive capabilities were obtained for deceleration patterns and 10-min Apgar score when evaluating CTG traces recorded on the day of delivery. The obtained results suggest the necessity of applying the criteria for the CTG traces

evaluation that are related to the gestational age. Our study corroborate the benefit of the CTG monitoring as a good screening procedure providing appropriate confirmation of fetal wellbeing.

ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Sciences and Higher Education resources in 2010-2012 under Research Project N N518 335935.

BIBLIOGRAPHY

[1] AYRES-DE-CAMPOS D., COSTA-SANTOS C., BERNARDES J., Prediction of neonatal state by computer analysis of fetal heart rate tracings: the antepartum arm of the SisPorto multicentre validation study, European Journal of Obstetrics & Gynecology and Reproductive Biology, Vol. 118, No. 1, 2005, pp. 52–60.

[2] BAMBER D., The area above the ordinal dominance graph and the area below the receiver operating characteristic graph, Journal of Mathematical Psychology, Vol. 12, No. 4, 1975, pp. 387–415.

[3] CZABAŃSKI R., JEŻEWSKI M., WRÓBEL J., JEŻEWSKI J., HOROBA K., Fuzzy system for evaluation of fetal heart rate signals using FIGO criteria, Journal of Medical Informatics and Technologies, Vol. 13, 2009, pp. 189–194.

[4] CZABAŃSKI R., JEŻEWSKI M., WRÓBEL J., KUPKA T., ŁĘSKI J., JEŻEWSKI J., The prediction of the low fetal birth weight based on quantitative description of cardiotocographic signals, Journal of Medical Informatics and Technologies, Vol. 12, 2008, pp. 97–102.

[5] FIGO News: Guidelines for the use of fetal monitoring, International Journal of Gynecology & Obstetrics, Vol. 25, 1987, pp. 159–167.

[6] JEZEWSKI J., WROBEL J., HOROBA K., KUPKA T., MATONIA A., Centralised fetal monitoring system with hardware-based data flow control, Proceedings of III International Conf. MEDSIP, Glasgow, 2006, pp. 51–54.

[7] JEŻEWSKI M., CZABAŃSKI R., HOROBA K., WRÓBEL J., ŁĘSKI J., JEŻEWSKI J., Influence of gestational age on neural networks interpretation of fetal monitoring signals, Journal of Medical Informatics and Technologies, Vol. 12, 2008, pp. 137–142.

[8] SIKORA J., Digital analysis of cardiotocographic traces for clinical fetal outcome prediction (in Polish), Perinatology and Gynecology, Vol. 21, 2001, pp. 57–88.

[9] STREET P., DAWES G.S., MOULDEN M., REDMAN C.W., Short-term variation in abnormal antenatal fetal rate records, American Journal of Obstetrics & Gynecology, Vol. 165, No. 3, 1991, pp. 515–523.