



Software for computerised analysis of cardiocotographic traces

M. Romano^a, P. Bifulco^b, M. Ruffo^b, G. Improta^b, F. Clemente^c,
M. Cesarelli^{b,*}

^a DMSC, University “Magna Graecia”, Catanzaro, Italy

^b DIETI, University of Naples, “Federico II”, Naples, Italy

^c IBB, Italian National Research Council, Rome, Italy

ARTICLE INFO

Article history:

Received 9 February 2015

Received in revised form

11 September 2015

Accepted 14 October 2015

Keywords:

Computerised cardiocotography

Foetal heart rate variability

Nonlinear indices

FHR clinical analysis

Software performance

Matlab

ABSTRACT

Despite the widespread use of cardiocotography in foetal monitoring, the evaluation of foetal status suffers from a considerable inter and intra-observer variability. In order to overcome the main limitations of visual cardiocotographic assessment, computerised methods to analyse cardiocotographic recordings have been recently developed. In this study, a new software for automated analysis of foetal heart rate is presented. It allows an automatic procedure for measuring the most relevant parameters derivable from cardiocotographic traces. Simulated and real cardiocotographic traces were analysed to test software reliability. In artificial traces, we simulated a set number of events (accelerations, decelerations and contractions) to be recognised. In the case of real signals, instead, results of the computerised analysis were compared with the visual assessment performed by 18 expert clinicians and three performance indexes were computed to gain information about performances of the proposed software. The software showed preliminary performance we judged satisfactory in that the results matched completely the requirements, as proved by tests on artificial signals in which all simulated events were detected from the software. Performance indexes computed in comparison with obstetricians' evaluations are, on the contrary, not so satisfactory; in fact they led to obtain the following values of the statistical parameters: sensitivity equal to 93%, positive predictive value equal to 82% and accuracy equal to 77%. Very probably this arises from the high variability of trace annotation carried out by clinicians.

© 2015 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

At the end of pregnancy, assessing of foetal well-being depends on the evaluation of multiple parameters, many of them correlated to characteristics of foetal heart rate (FHR) signals recorded by means of Cardiocotography (CTG). This technique, widely used both in the antepartum and

intrapartum period [1], consists of the simultaneous recording of FHR and uterine activity (generally referred as uterine contractions signal – UC) [2,3]. Both FHR and UC are simultaneously recorded and printed onto a paper strip. In the classical approach, clinicians, whose task is to classify the signal patterns, visually interpret this graphical representation. Many parameters of FHR signals need to be interpreted for the evaluation of foetal health status: the basal level of FHR signal

* Corresponding author. Tel.: +39 081 7683788; fax: +39 081 7683804.

E-mail address: cesarelli@unina.it (M. Cesarelli).

<http://dx.doi.org/10.1016/j.cmpb.2015.10.008>

0169-2607/© 2015 Elsevier Ireland Ltd. All rights reserved.

(also called baseline), FHR variability (FHRV) and transient increases (accelerations) or decreases (decelerations) of the FHR. FHRV represents the beat-to-beat effect of foetal sympathetic and parasympathetic nervous systems and its changes may be also the result of the foetal sleep-wake cycle, drugs, umbilical cord compression, foetal hypoxia and such severe conditions as foetal acidemia [4]. Accelerations are the result of foetal movements and identify the foetal well-being, while decelerations are the symptom of foetal distress usually indicating the risk of foetal hypoxia [5], even if their interpretation is difficult and associated to different aspects such as shape, superimposed variability and delay with respect to UC [6].

Visual analysis of CTG recordings (sometimes here called CTG traces or only CTG) has a well-demonstrated poor reproducibility [7] due to the complexity of physiological phenomena affecting foetal heart rhythm [8] and being related to clinician's experience. A sufficiently high level of disagreement was noted between observers as well as for the same observer [9–11]. This great intra- and inter-observer variation negatively affects sensitivity and specificity in reading CTG traces [12]. Besides, abnormal patterns can be recorded both in case of foetal distress and in foetal well-being but during quiet periods of the foetus. Therefore, establishing foetal state is often a difficult issue in clinical practice and the erroneous evaluation of CTG may lead to unjustified invasive medical interventions. This is not just an ethical issue since in some countries CTG recording is a medical report with legal value [13–15].

In order to decrease the subjective nature of foetal state evaluation, the visual interpretation is more and more often replaced by automated computerised analysis, which has the theoretical advantage of providing a reproducible and objective interpretation of CTG traces and quantifying parameters that are difficult to assess by the human eye, such as short- and long-term variability [7,16]. Furthermore, it has been demonstrated that computerised CTG has also a great potential to improve the diagnosis of foetal asphyxia in labour [17]. Computerised systems are also able to provide an easily accessible means for storage, review and transmission of data and enhance the constitution of databases, with important clinical and research applications [7]. They have been developed since the late 1970s but the interest in this field is still alive since computerised CTG systems have not yet gained very wide clinical acceptance. This may be related to the limited demonstration of their validity and efficacy, as well as to the poor practicality of their routine clinical use [16]. At the same time, it is difficult to implement the heuristic rules used by clinicians during the diagnostic inference in signal processing algorithms [5]. Up to now, the basic idea in automatic CTG analysis has been the assessing of parameters historically and habitually used by clinicians for the diagnosis such as the baseline, followed by the detection of events like accelerations and decelerations, considering and measuring some morphological characteristics in the same way as the clinicians do by eye inspection [18]. Nevertheless, this is not a simple task; in particular, there is currently no consensus on how computer estimation of the FHR baseline should be performed [19], although the baseline is considered as one of the fundamental features of the FHR pattern because it can influence overall tracing interpretation and most of the other features rely on

its estimation [20,21]. Besides, there is not still a unique criterion for defining accelerations, and analogous disagreement between authors can be found in establishing criteria for the detection of decelerations and contractions. Thus, developing a software that can be really helpful for clinicians in foetal monitoring represents a remarkable purpose.

The aim of this study is to present the main characteristics and potentiality of a new software for computerised analysis of antenatal CTG developed mainly for research purposes, which, although does not represent a diagnostic tool, can be also a useful technical support for clinicians. It has been the result of a research study of some previous automated systems for FHR analysis present in literature in order to join the estimation of the most relevant parameters for a more objective analysis. First target of the proposed computerised analysis system is to make automatic the procedure for measuring the parameters which are historically and still nowadays employed in daily clinical practice. Then, since many works highlight the usefulness of nonlinear techniques in the study of heart rhythm [22–25], other less traditional parameters are involved in order to enlarge information amount related to FHR.

2. Materials and methods

In this section an overall description of CTG recordings, software features and tests carried out is given; a more detailed analysis is reported in [Appendix A](#). The software was completely developed in Matlab ver. R2011a. Concerning classical parameters, the implementation of the algorithm started from the indications by Mantel et al. [26,27], because they seem to be the most complete and detailed.

2.1. Data set

CTG traces here used to test software performances are part of a database which includes almost 400 signals recorded from 2000 to 2009 in clinical environment. All CTG were recorded during daily routine foetal monitoring in clinical environments from women between 31st and 41st week of gestation, both in antepartum and in intrapartum period; they last at least 20 min. The patients laid down in a rest position. Neonatal parameters such as Apgar score and new-born weight and other information such as eventual maternal pathologies had been also collected when available. Based on the available data, about 55% of pregnant women had a caesarean section and 35% a spontaneous delivery. In addition, 45% of births were male and 35% female.

Cardiotocographic signals were acquired using HP-135x or Sonicaid cardiotocographs equipped with an ultrasound Doppler probe to detect FHR signal and an external pressure transducer to record UC signal. In HP cardiotocographs, FHR and UC signals are internally stored at 4 Hz (corresponding to a sampling interval of 250 ms). On the contrary, in Sonicaid cardiotocographs, FHR and UC signals are unevenly stored.

Digital signals were acquired from the CTG devices to the computer's serial port via a standard computer cable, using a RS232 protocol. An interface (not described for sake of brevity)

was developed by the authors in C language only for CTG acquisition, storing in files and real-time visualisation.

For considering at least 30 signals, initially, 40 CTG were randomly extracted from that database. Then, in order to involve only CTG signals by healthy foetuses and of acceptable quality, were excluded from the study signals that did not meet all the following requirements:

- a length of at least 20 minutes;
- a percentage of signal loss lower than 10%;
- a percentage of outliers lower than 5% (we judged a signal with too many outliers poorly reliable);
- acquired from singleton pregnancy;
- spontaneous delivery or elective caesarean, without complications;
- correspondence with a healthy newborn (Apgar at least 7, birth weight appropriate for gestational age).

Some of these informations, such as percentage of signal loss and outliers, are provided by the pre-processing (Section 2.2).

After this check, 34 CTG signals, with different levels of interpretation difficulty, were considered in the study.

2.2. Pre-processing

This section of the software is not object of our tests and is briefly described only for sake of completeness, its functioning and performances were previously reported [14,28]. It is blind to the operator and has as main purpose to make available for further analysis FHR and UC signals with similar characteristics regardless from the device used for CTG acquisition (the algorithm can process CTG signals recorded with HP/Philips, Sonicaid or Coronometrics devices).

The software devoted to the pre-processing processes files in output from the interface. Main steps of the algorithm are:

- signal reading and assignation of the variables: FHR, toco (uterine contractions) and quality;
- recognition of the signal tracts having good and bad quality (these last including tracts of signal loss).

For each segment of good quality:

- recovery of the real uneven FHR series when CTG output is evenly spaced (case of HP/Philips cardiotocographs) [14];
- processing of outliers [28]:
 - detection and substitution of isolated outliers;
 - detection of consecutive outliers, whose quality is posed equal to zero.
- interpolation of signal tracts of poor quality (according to an index provided by the equipment) or signal loss which last maximum 3 s, in order to avoid an excessive fragmentation of the signal.

These operations, actually necessary only for FHR signals, were simultaneously adopted also for UC signals in order to maintain the alignment between them.

Besides, a time axis in seconds is constructed summing from time to time the reciprocal of FHR samples, expressed in beats per minute (bpm), multiplied by 60.

In output, the algorithm provides also other information useful to increase the reliability of successive processing, such as quality of FHR samples, duration of the more length segment of good quality, position of samples corresponding to substituted outliers, percentage of remaining outliers, position and percentage of interpolated samples and percentage of loss or bad quality tracts in the output FHR signal.

CTG traces populating the data set described in Section 2.1 are output of the pre-processing.

2.3. Classical analysis

This part of the analysis is the core of our software since it provides the estimation of the main clinical-interest parameters in foetal monitoring [29].

About this analysis, it is worth highlighting the absence of widely accepted definitions and/or quantitative criteria to be directly implemented in computerised CTG. For example, concerning accelerations, events more easily recognisable, according to Mantel et al. [27], an acceleration should have a peak >10 bpm above the baseline and should last for more than 10 s. FIGO guidelines [30] and the National Institute of Child Health and Human Development (NICHD) [31], instead, proposed a definition of acceleration like an increase of 15 bpm from the baseline with a duration longer than 15 s. Other examples could be reported, such as the introduction of small and prolonged accelerations [31,32]. Even about the definition of the normal range of FHR there is still an open debate [33].

In this scenario, our algorithm for the identification of classical parameters is based on a mixture of rules based on FIGO guidelines, other literature indications [2,13,26,27,31,34] and experimental choices (please see Appendix A for details).

2.3.1. Baseline estimation

Following Mantel criteria and FIGO guidelines, the baseline is defined as current average of the heart rate in absence of accelerations and decelerations [26,30]. In particular, the baseline is computed in five steps. At first, a low pass filter on the FHR signal is used in order to obtain a starting baseline estimation. Secondly, a lower (20 bpm below the previous estimated baseline) and a higher (20 bpm above the baseline) threshold is set to remove FHR variations that exceed this range and a low pass filter is adopted afterwards for evaluating the baseline of the new FHR. This second step is reiterated three times changing only the higher threshold value, which is reduced by 5 bpm for every iteration.

2.3.2. Accelerations detection

The algorithm for the identification of accelerations provides the total number of accelerations, their classification (big, small or very small) and, on demand, all their numerical characteristics (start and end, duration, maximum amplitude, area).

It is divided into three main phases. First, preliminary identification of candidate accelerations (based on criteria concerning maximum amplitude with respect to the baseline and duration). Secondly, control of the gap (lowering of

the heart rate below a certain threshold within the same acceleration) and, in case, division in two or more candidate accelerations, which could be not considered in the subsequent steps if they do not satisfy fixed rules. Finally, elimination of accelerations containing too many interpolated values within them (on the basis of an accepted maximum value heuristically set).

2.3.3. Decelerations detection

The algorithm for the identification of the deceleration is similar to that of the accelerations. It has the same logical flowchart but, of course, adopted criteria are different for all the phases.

2.3.4. Floatingline evaluation

On the basis of a previous work of the authors [35], the floatingline estimation was implemented by means of a nonlinear filtering, using the cubic smoothing spline and varying the smoothing parameter (corresponding to different cut-off frequencies) to obtain a curve that comes closest to the definition of floatingline, median line of the FHR. In particular, the filter band is different in presence or absence of accelerations (which enlarge the frequency band of the floatingline). After that, just for visual inspection, the various tracts of bad quality are linearly interpolated.

2.3.5. Contractions assessment

With regard to the detection of uterine contractions, a reference is calculated (a kind of basal tone) with an algorithm very similar to that developed for the calculation of the baseline; then, uterine contractions are detected by evaluating their duration and amplitude with respect to the computed basal tone.

2.3.6. Other parameters

In addition to the features described above, the software provides the value of FHR mean, the length in minutes of the signal and the total percentage of interpolated samples.

2.3.7. CTG visualisation

With regard to the graphics, in addition to the typical display of Matlab (default visualisation), the user can choose to view the CTG trace according to clinical standards. In this case, the software simulates the CTG paper running speed of 1 centimetre (cm) per minute and displays trace segments of 25 cm for every page. The FHR range is 50–210 bpm and UC range is 0–100 a.u.

In order to distinguish the different features, we choose various colours for each of them: FHR and UC signals are printed in black, baseline in magenta, floatingline in yellow, accelerations in green, decelerations in blue, contractions in cyan.

In addition, accelerations lasting from 3 to 10 min, or decelerations lasting from 2 to 3 min, i.e. those that do not meet FIGO definitions, are highlighted with a message and green (or blue for decelerations), black-bordered triangles. Finally, it is possible to point out signal tracts containing interpolated values, highlighted with red crosses. Also in this case, if the number of continuous interpolated samples exceeds a certain value, a legend warns the clinician about their presence.

2.4. Other analyses (time domain, frequency domain, nonlinear techniques)

In order to complete the analysis and to enlarge the data set of CTG features to be supplied to the user, the software estimates other FHR parameters obtained with different techniques: Time and Frequency Domain analyses, recognised as important tools for assessment of foetal well-being, and the more recent Symbolic Dynamics.

2.4.1. Time domain: STV index

Short term variability (STV) of FHR refers to the continuous variation in difference between successive inter-beat intervals and it is difficult to interpret reliably with the naked eye. This parameter is important since permits to investigate foetal reactions to internal or external stimuli. In general, large variability reflects a healthy autonomic nervous system (ANS) and, for this reason, STV can represent a valid support to diagnose foetal health.

In agreement with a previous study of the authors, the software assesses STV as standard deviation of the FHRV, obtained after subtraction of the floatingline from the FHR signal [36]. The computation is carried out on sliding windows of length M (with M covering 30 s), with an overlap of $M - 1$ samples; then, for providing an overall STV index of the signal, the mean of all STV values is computed.

2.4.2. Frequency domain: short time Fourier transform

According to the literature and methodologies previously used by the authors [37–40], the power spectral density (PSD) of the FHRV is estimated by means of the Short Time Fourier Transform (STFT) using a sliding Hamming window of length L (where L corresponds to 32 s), with an overlap of $L - 1$ samples, after removal of the floatingline [41]. For each shift of the window, the software computes a new PSD. To be able to use the STFT, the FHR signal is previously interpolated by means of a cubic interpolation and re-sampled at 4 Hz [42]. The main frequency bands of FHR, chosen after a careful study of different literature works [43–46], are set from 0 Hz up to 0.05 Hz for very low frequencies (VLF); from 0.05 Hz up to 0.2 Hz for low frequencies (LF); from 0.2 Hz up to 1 Hz for high frequencies (HF). Then, the software computes the power in these bands (respectively named VLFp, LFP and HFP) by means of a simple integral rule and the sympato-vagal balance index (SVB), as ratio between LFP and HFP.

2.4.3. Frequency domain: Lomb

Although it has been demonstrated that the cubic interpolation minimises the errors introduced in the PSD, in order to avoid this operation, the Lomb method is also employed to compute the PSD [47–50]. This method, in fact, was suggested for HR signals in adult subjects by Laguna et al., just to be applied directly on the uneven series; even if it has the disadvantage to be very time consuming. For a more detailed description of the technique, please refer to specific literature or to previous works of the authors [41,42]. After the PSD evaluation, the same frequency parameters are computed (VLFp, LFP, HFP, SVB).

2.4.4. Nonlinear techniques: symbolic parameters

The Symbolic Dynamics analysis (SDA) is a nonlinear technique here introduced since in the last decades it has been used in adult HRV analysis for its potentiality in highlighting nonlinear mechanisms involved in heart rhythm control and it has already obtained some success [51,52]. Although in a less widespread, it has been employed also in foetal field providing promising results [22], which indicate that it could represent a helpful additional tool in foetal monitoring.

For providing a very brief description, main steps of SDA are: encoding of a series using symbols chosen by a previously set alphabet and defined rules, generation of words of fixed, adequate length and classification of the words.

Recently, the authors have proposed an index, computed analysing the distributions of words with similar characteristics, as an alternative and complementary index to estimate the variability of FHR signals and to study the relationship between FHRV and neonatal wellbeing [53–55]. This index, named V.I. (Variability Index), is included in the set of parameters provided by the software.

2.5. Graphic interface

A graphic interface, enforced with Matlab Graphical User Interface Development Environment (GUIDE), allows clinicians to interact with the software by means of simple buttons. The different analysis, classical, time domain, frequency domain and nonlinear are independent.

Besides, the operator can modify the baseline smoothing value (of course a default value is set). When the user chooses this option, he is warned that, as consequence, results of classical analysis will change and, in any case, changes will apply only for the current session.

The results of the analysis by default are shown at monitor but, of course, can be also printed and/or saved.

2.6. Simulated CTG signals

In order to verify the agreement between software results and specifications, we carried out tests on synthetic CTG signals.

According to previous works, an artificial R–R tachogram with R–R samples expressed in seconds and with specific power spectrum characteristics, proper for FHR, was generated using an inverse Fourier transform. In particular, we fixed central frequencies of the spectral bands, the bandwidths and the ratio between power of low and high frequency bands. Then, the FHR signal was computed using the known formula $FHR = 60/RR$, and fixing FHR mean value and standard deviation of its peak-to-peak amplitude. After that, accelerations and decelerations were simulated by using Gaussian-like signal tracts and setting a priori waveform's features and parameters, such as for example amplitude, position and duration (for a more detailed explanation about how these synthetic signals were generated, please refer to previous works of the authors [14,36,56]).

UC signals were simulated directly in time domain. The basal UC signals were simulated by low-pass filtering a white noise and, as for the FHR, a sinusoidal basal tone can be superimposed. Then, it is possible to simulate the contraction

events with a procedure similar to that used for the accelerations.

In this way we knew exactly how many and what events were present, i.e. we had a gold standard with which compare software results, hence we could evaluate its performances.

2.7. Clinical validation

In case of real signals, because of the lack of a gold standard to be used for the validation of the software, we decided to compare the results of Section 3.5 with the visual assessment performed by a team of 18 expert clinicians who referred mainly to FIGO guidelines. All them had more than 10 years experience in CTG analysis and continue to use this technique during their daily clinical activities almost all in the most important hospital of South Italy. In order to avoid a possible source of error, CTG signals printed in way to be compatible with the usual clinical format (please, see Section 2.3.7) were submitted to their analysis. Of course obstetricians did not know results of computerised analysis.

This kind of comparative analysis is not unusual, in fact, traditional CTG inspection is widely considered an important basis to test the reliability of computerised systems [4,19–21,32,57,58], even though the subjectivity of visual assessment and the lack of an objective measure of foetal condition cause an unavoidable inter-observer variability [9,27].

The comparison was conducted submitting all 34 CTG traces to be analysed, to all experts and to software.

We focused our attention specifically on events detection (accelerations, decelerations, contractions). Basically, we compared the number of the events recognised through our software with the ones identified by means of visual inspection, regardless of the instants of beginning and end, and then we computed three indexes representing measures of our software technical performance (for details see following Section 2.8).

We neglected baseline and floatingline testing because despite to their importance they generally are not directly estimated by clinicians, even if baseline is the reference used to detect accelerations and decelerations.

The other analyses were excluded from the tests because in most cases they can be executed only in numeric way and some of them are not so familiar for clinicians.

2.8. Event identification

After an initial, qualitative evaluation of the comparison results, we noted that many divergences between obstetricians and software results were due to a different approach in some situations.

Therefore, as consequence of some discussions with clinicians and a deepened analysis of the events, we decided to exclude from the study the following categories of events:

- events occurring at the initial or final tract (length heuristically set at 1min) of CTG signals were neglected by clinicians;
- in case of two or more consecutive, very close or overlapped accelerations (distance of about 6s or lower), only one event is usually considered by clinicians whereas the other ones

are not taken into account because they do not provide additional information to the recognised foetal activity;

- still concerning the accelerations, those classified by the software as “non FIGO” were excluded because of unclear definition;
- about UC signals, some obstetricians recognised also spikes, identified as foetal movements and hence considered of prognostic meaning. Nevertheless, these events (with duration lower than 20 s) were did not involved in the study since they do not meet literature definition of contractions neither of foetal movement (at the best of our knowledge a definition of foetal movement is completely missing and entrusted only to the physician experience).

In the following Fig. 1, some examples of excluded events are shown.

In the above illustrated CTG trace, it is possible to observe an example of two consecutive FHR accelerations between the 4th and the 5th min of recording with an estimated distance of about 3 s, which are well identified by the software but, at the same time, are treated as only one overall event by clinicians inspecting the signal. Moreover, within the same recording, it is clearly displayed a spike at about the 13th min of UC signal, which obstetricians considered as a foetal movement. Both those events were excluded from the study.

2.9. Performance indexes estimation

Due to the high inter-observer variability (found also in this study, as reported in Section 3), different methodologies there exist in recent literature to specifically evaluate agreement computer-observer or among observers [59–61]. Nevertheless, there is no a reference gold standard to establish the true positive events.

In this study, we defined a weight factor (W in Eq. (1)) representing the probability that an event is true in accordance with the number of obstetricians who recognised it.

Then, we set two thresholds on the basis of two percentages of agreement between clinicians who examined the recordings, Perc1 and Perc2, respectively equal to 15% and 85% in this study. We considered $W = 1$ if the percentage of agreement is at least Perc2 and $W = 0$ for a percentage lower than Perc1. Between Perc1 and Perc2, W is linearly varying.

$$W = \frac{N - \text{round}(\text{Perc1} \times \text{NG})}{\text{round}(\text{Perc2} \times \text{NG}) - \text{round}(\text{Perc1} \times \text{NG})} \quad (1)$$

where NG = total number of clinicians involved in the study; N = number of clinicians who identified a specific event; Perc1 = percentage of identification leading to consider $W = 0$; Perc2 = percentage of identification leading to consider $W = 1$.

In order to have an index limited between 0 and 1, we set at 0 all negative results and at 1 all W values greater than 1.

Finally, we defined three performances indexes associated to the events classified as true positive, false positive and false negative (as known, events not detected by the software),

Table 1 – Results of tests carried out using simulated FHR signals.

	Simulated events	Recognised events	Percentage of identification
Accelerations	120	120	100%
Decelerations	60	60	100%
Contractions	30	30	100%

respectively named IP_{TP} , IP_{FP} and IP_{FN} , computed according to the following formula:

$$\begin{aligned} IP_{TP} &= \sum E \times W \\ IP_{FN} &= \sum \bar{E} \times W \\ E &= \sum E \times (1 - W) \end{aligned} \quad (2)$$

where IP = Index of Performance; $E = 1$ if the software identifies the event, 0 otherwise; \bar{E} = complement of E ; W = weight factor; TP = true positive; FN = false negative; FP = false positive

2.10. Statistical parameters

As last step, we evaluated some statistical parameters in order to verify the overall potentiality of the software.

We computed sensitivity, positive predictive value and accuracy (see formula below) substituting numbers of true positives, false negatives and false positives by the corresponding performance indexes.

$$\text{SENSITIVITY} = \frac{\text{true positive}}{\text{true positives} + \text{false negatives}} \times 100$$

POSITIVE PREDICTIVE VALUE

$$= \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \times 100 \quad (3)$$

ACCURACY

$$= \frac{\text{true positives}}{\text{true positives} + \text{false positives} + \text{false negatives}} \times 100$$

3. Results

3.1. CTG visualisation

Fig. 2 shows an example of CTG visualised according to clinical standard, with the detected events highlighted by different colours.

3.2. Tests on simulated CTG signals

Table 1 shows results and percentages of detected events in simulated FHR signals.

3.3. Clinical validation and performance indexes

Results concerning the clinical validation initially were not satisfying, so that a new meeting for defining events inclusion criteria in a more and more precise way and sharing observations with clinicians was necessary. Mainly, it was clarified

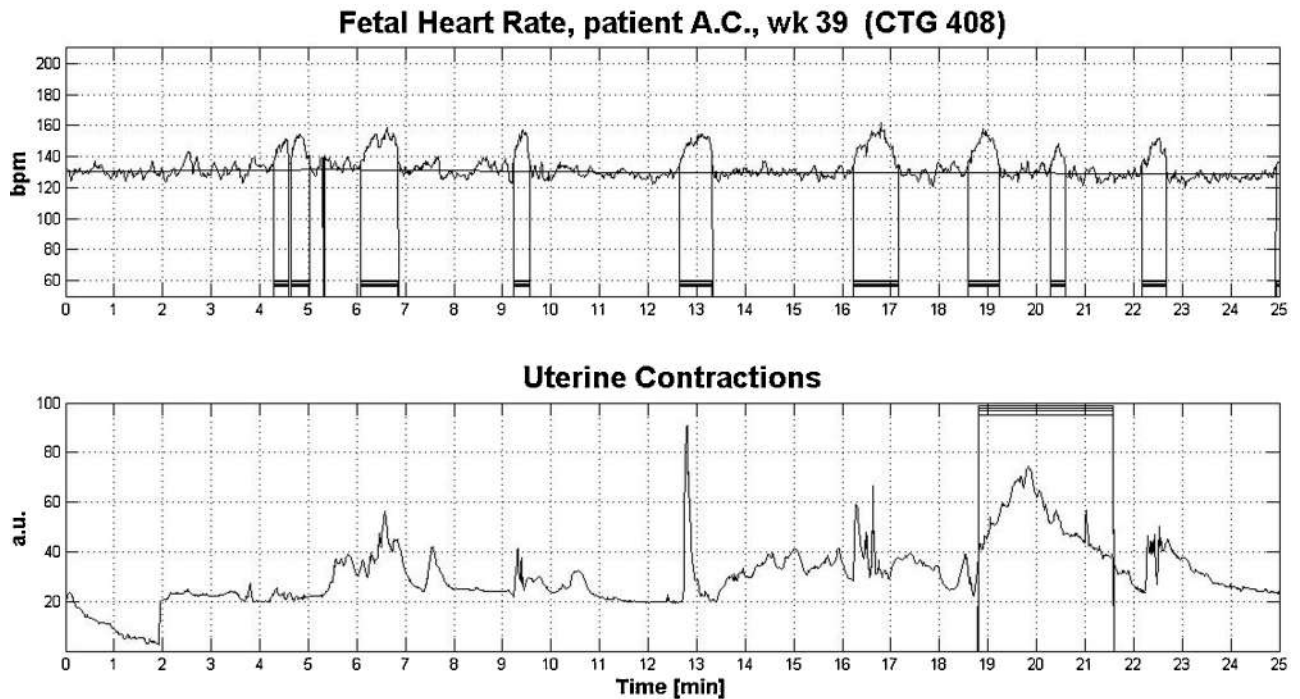


Fig. 1 – Trace CTG408 (internal number of our database). Between the fourth and fifth minute of the FHR signal, an example of too near accelerations not involved in the study. Approximately at the thirteenth minute of the UC signal an example of spike not considered as uterine contraction.

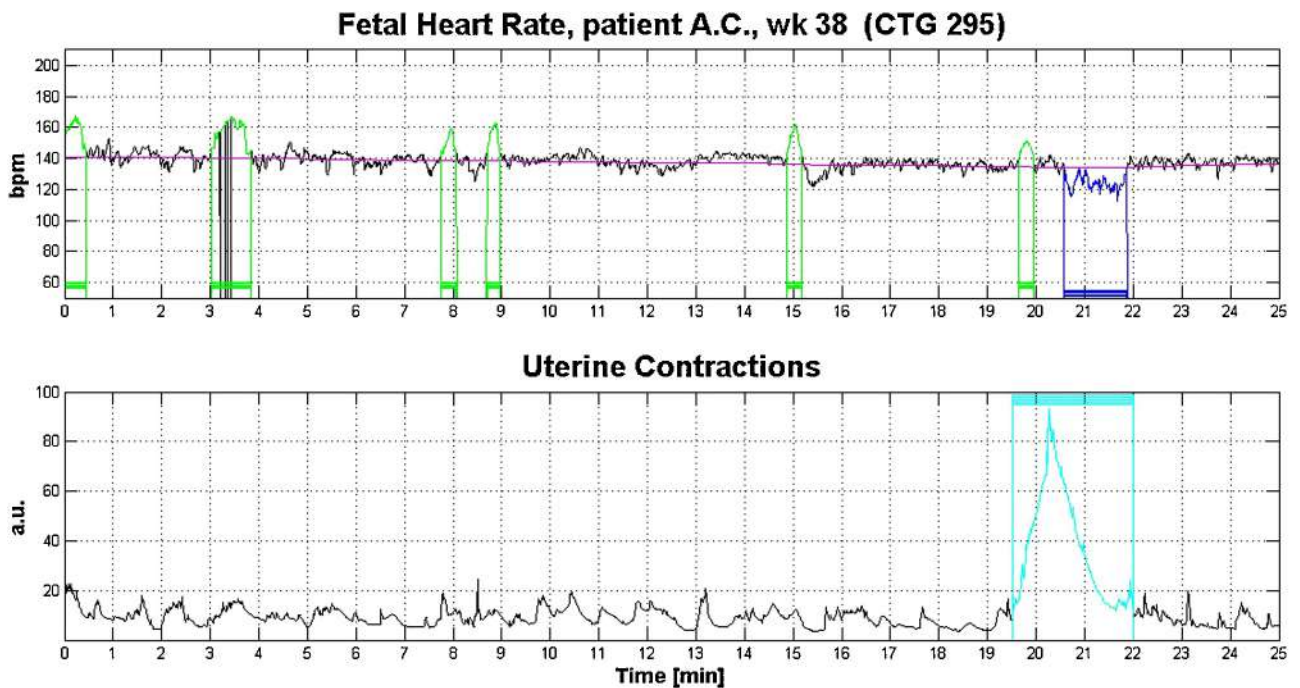


Fig. 2 – Example of visualised CTG signal. FHR on the top, UC on the bottom.

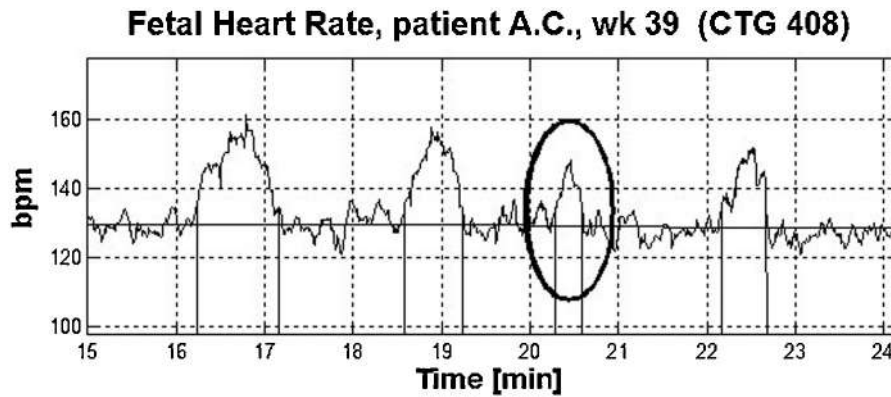


Fig. 3 – In the oval an example of acceleration (CTG #408) included in the evaluation by clinicians only after the second meeting.

Table 2 – For each event considered, % tot represents the percentage of events for which all obstetricians were in agreement and % 50 is the percentage of events recognised by at least 50% of the obstetricians.

	Total events	% tot	% 50
Accelerations	296	36.5	85.8
Decelerations	18	11.1	61.1
Contractions	70	54.3	71.4

Table 3 – Values of performance indexes.

	IP _{TP}	IP _{FP}	IP _{FN}
Accelerations	223.08	45.92	15.42
Decelerations	7.25	1.75	1.92
Contractions	45.33	14.67	4.42

that also events poorly meaningful in the contest of the whole FHR signal had to be considered in the study if matching the definitions. In Fig. 3, it is shown, as an example, an acceleration initially discarded because considered not important in the contest of the whole FHR signal (sufficiently variable and with the presence of many greater accelerations).

In Table 2 we reported, as percentage values, the final results of accordance among obstetricians for each kind of event here considered.

Values of performance indexes, computed in tests carried out on the data set of real CTG signals, are summarised in Table 3. It is possible to observe that due to the choice of introducing a weight factor, they are not integer values.

For sake of completeness, it is interesting to observe that the software does not detect only less than 7% of events recognised by at least 50% clinicians.

Finally, it is worth highlighting that events for which all obstetricians were in agreement very, very probably can be considered “true events” and that the software recognises all of them.

3.4. Statistical parameters

Finally, in Table 4 we reported the results of the statistical analysis.

Table 4 – Values of statistical parameters for each event and for the total number of events.

	Sensitivity	Positive predictive value	Accuracy
Accelerations	94%	83%	78%
Decelerations	79%	81%	66%
Contractions	91%	76%	71%
Overall	93%	82%	77%

Values obtained using simulated signals are not reported since all parameters resulted equal 100%.

3.5. Classical analysis

Here, we report some results of the software concerning the classical analysis of real CTG signals.

In the first example, Fig. 4, we report an example of CTG trace in which the software recognised all kinds of accelerations (very small, small and big), a deceleration and different contractions (small and big); for details about classification criteria of the events, refer to Appendix A.

In Fig. 5 an example of CTG traces with repetitive contractions and several accelerations and decelerations.

In this case, we used the default display of Matlab and not the clinical visualisation in order to show a longer signal tract.

In Fig. 6, instead, again with the default visualisation, an example of software performances in analysing a CTG trace characterised by high variability, for which it was still able to detect heart foetal reactions (accelerations) to movements and/or uterine activity.

3.6. Other techniques

Just for sake of completeness, in this subsection we reported some example of results of other techniques carried out by the software.

3.7. Time domain: STV

In addition to the index, at user request, the software provides a graphical representation of all STV values, computed in each

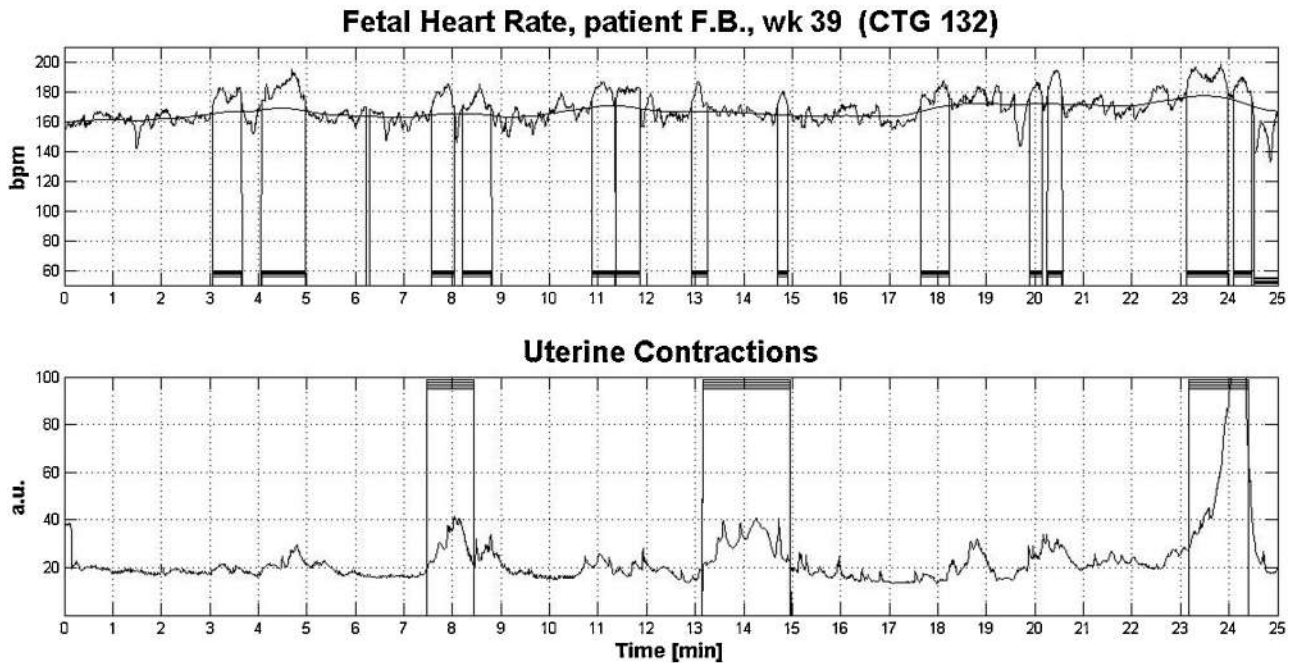


Fig. 4 – CTG trace in which the software recognised all kinds of accelerations, a deceleration and different contractions (small and large). For example, accelerations at around 15 and 20 min were labelled as very small and the first uterine contractions were considered small.

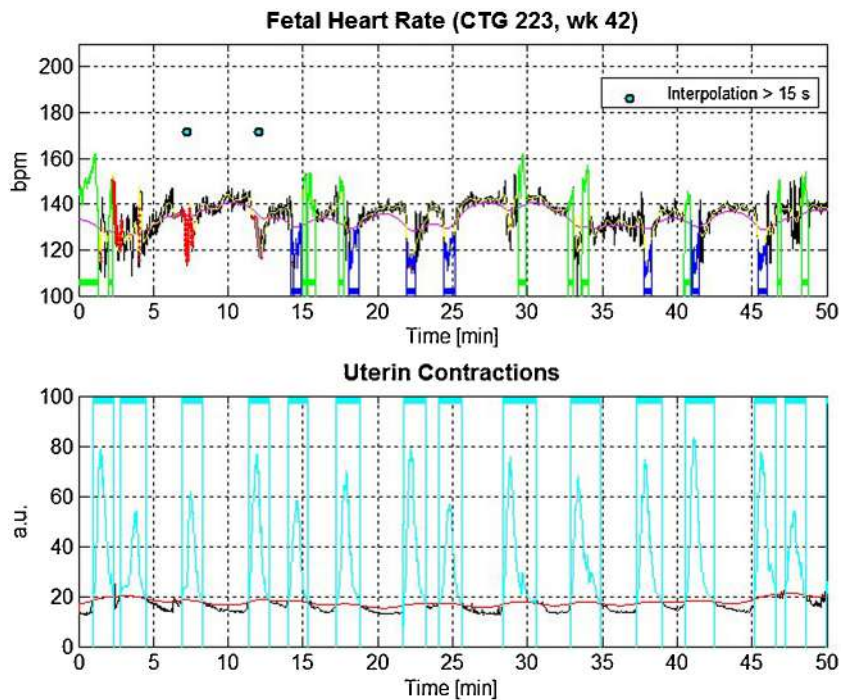


Fig. 5 – First part of analysis carried out for CTG #223 (internal numbering of our database). It is possible to see, over the analysis results, the warning in correspondence of interpolated samples (highlighted also by red markers). Here we used the default display of Matlab and not the clinical visualisation in order to show 50 min in a single figure.

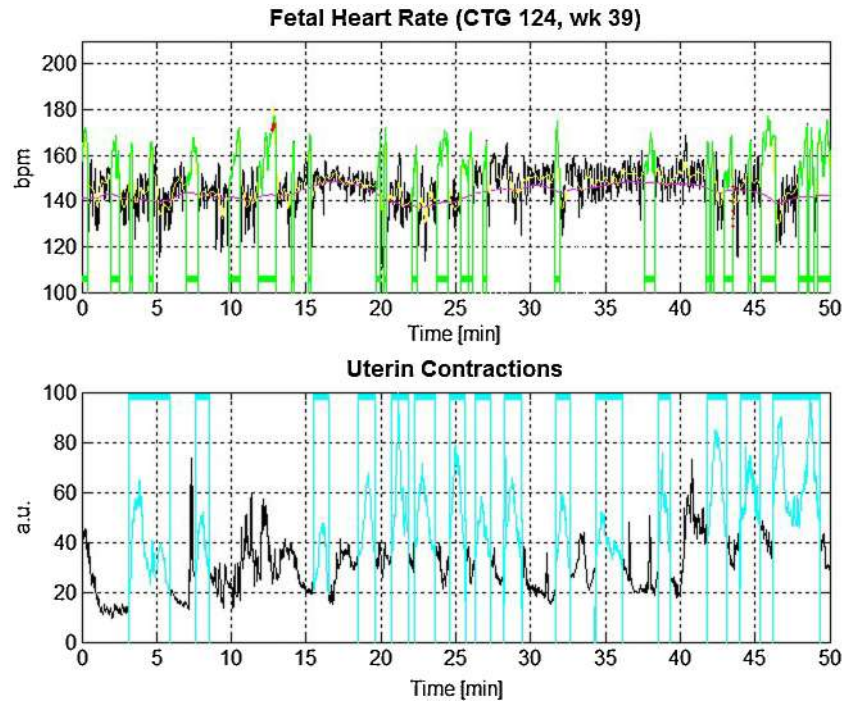


Fig. 6 – Analysis of the CTG #124.

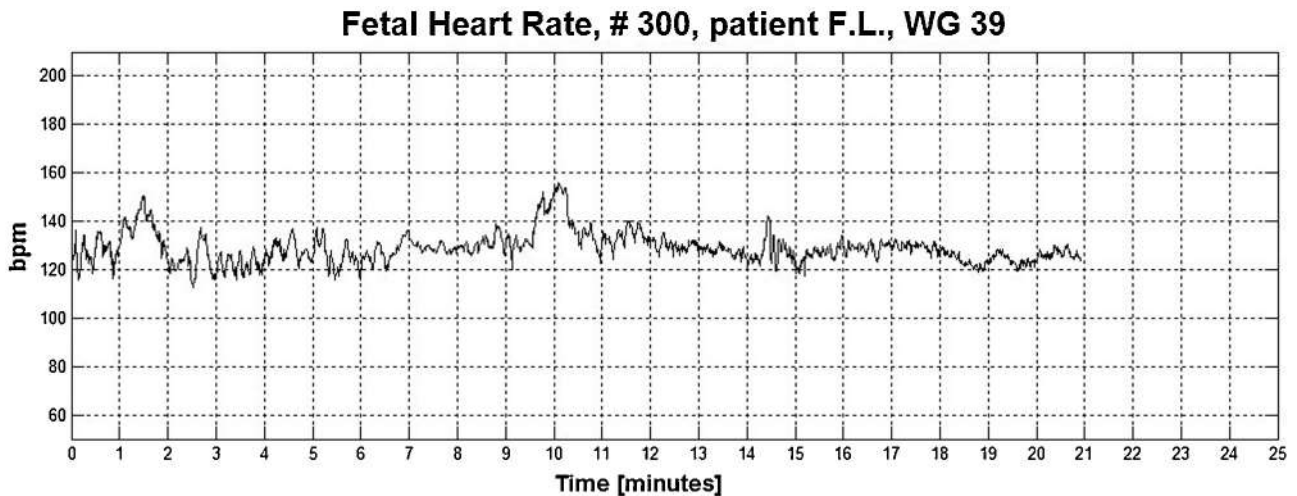


Fig. 7 – FHR 300 printed according to clinical standard before analysis.

window of 30s, along with the FHRV signal. In Fig. 8 is shown as example the STV trend of the signal depicted in Fig. 7 for which a STV value equal to 2.28 bpm was obtained.

3.8. Frequency domain: STFT

For each FHR signal, all PSD estimated are stored in a time-frequency matrix so that, if requested, the software can show the mean power spectrum and other parameters such as power content of the different frequency bands. In Fig. 9 we reported an example of mean power spectrum to which correspond the following values of power content: $LFp = 5.12 \text{ bpm}^2$ and $HFp = 2.42 \text{ bpm}^2$.

3.9. Symbolic dynamics analysis

In this case the software displays the histogram of word classes and the value of the V.I. (see Appendix A for details), as depicted in Fig. 10, for which we obtained $V.I. = 0.5$.

4. Discussion and conclusion

As it is known, computerised CTG analysis offers many advantages; among them, it is worth mentioning that predefined criteria to be satisfied are always met, obtaining a really more objective FHR reading, as demonstrated, for our software, by tests on simulated signals, satisfying in 100% cases of event

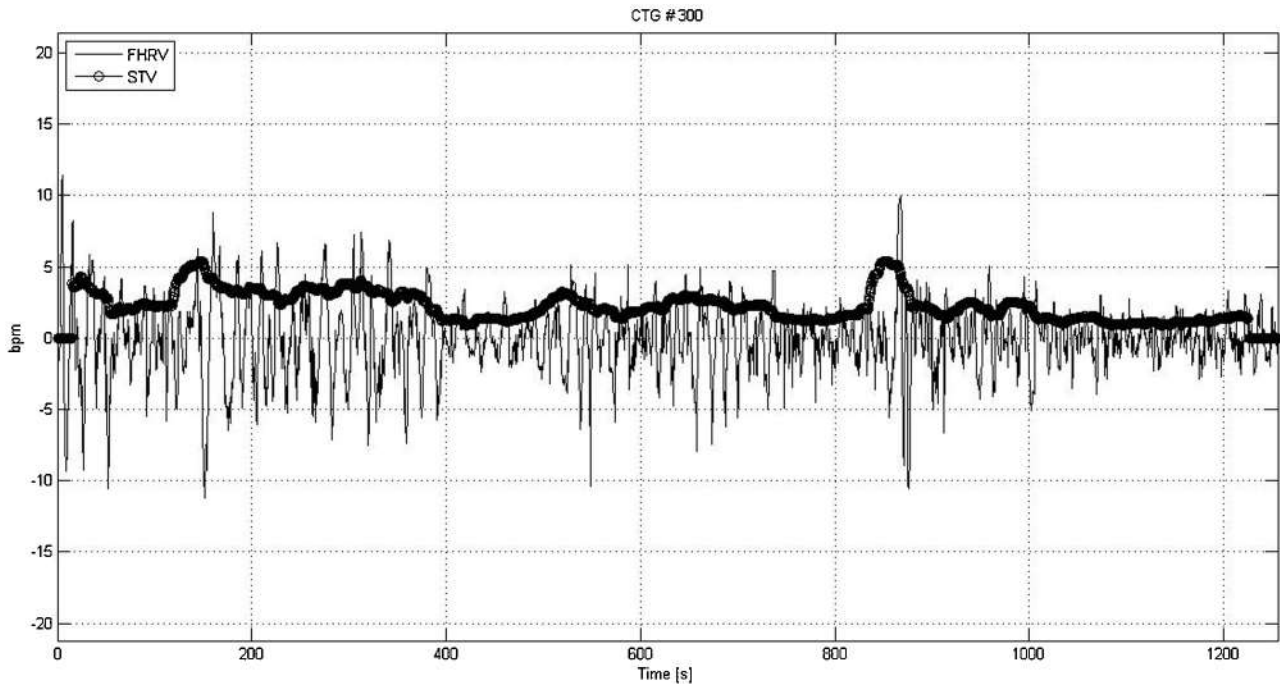


Fig. 8 – FHRV signal with superimposed, plotted with circles, STV time-course point-by-point. STV index = 2.28 bpm.

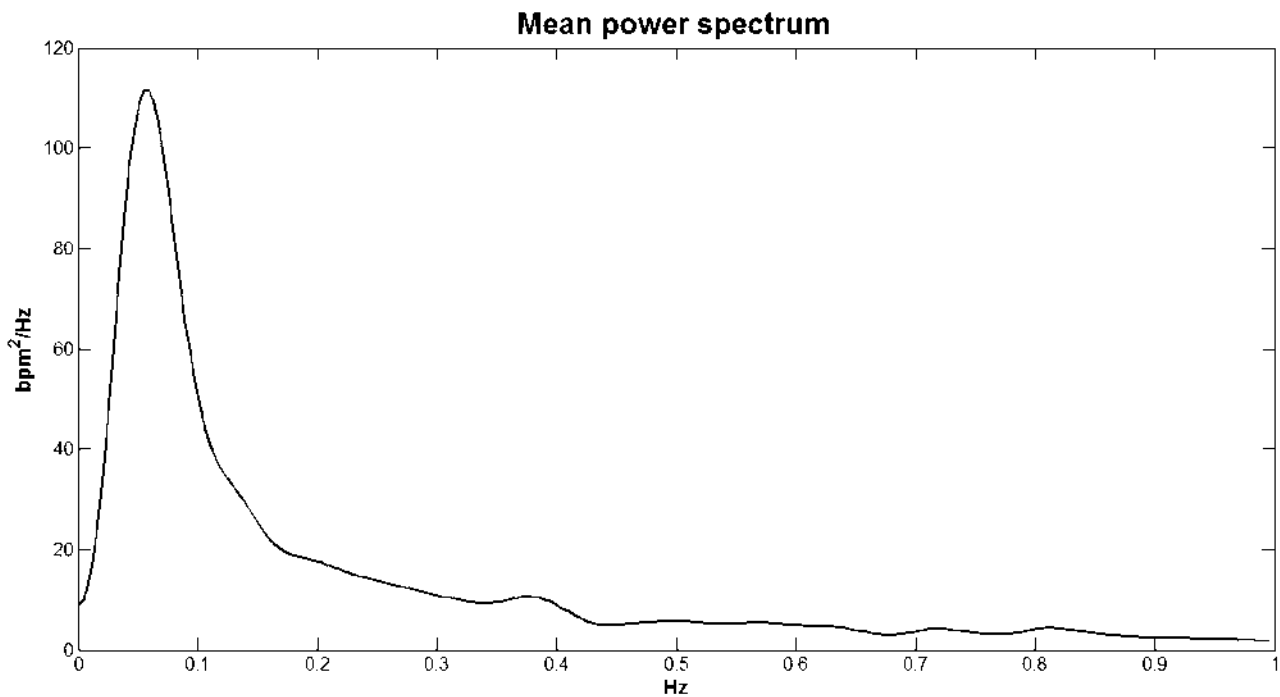


Fig. 9 – Mean power spectrum computed by means of STFT for the FHR 609. $LF_p = 5.12 \text{ bpm}^2$; $HF_p = 2.42 \text{ bpm}^2$.

identification. Further, quality of signals can be improved by an adequate pre-processing; computerised systems can be developed to be interactive and give warning signals in case of necessity. (We planned a similar feature for the on-line version of the software.)

Due to the increasing importance of computerised CTG, different interesting examples of software for automatic CTG analysis were described in literature, some of them also

commercially available. Although many aspects are common for all systems, each computerised CTG has relevant characteristics. For example, the Sonicaid Fetal Care is the only system based on the Dawes/Redman analysis criteria [34]. The SisPorto 2.0, latest version of a programme for automated analysis of CTG traces developed at the University of Porto, is able to acquire signals from whatever foetal monitors [7,62]. The 2CTG2 system involves in the analysis CTG parameters

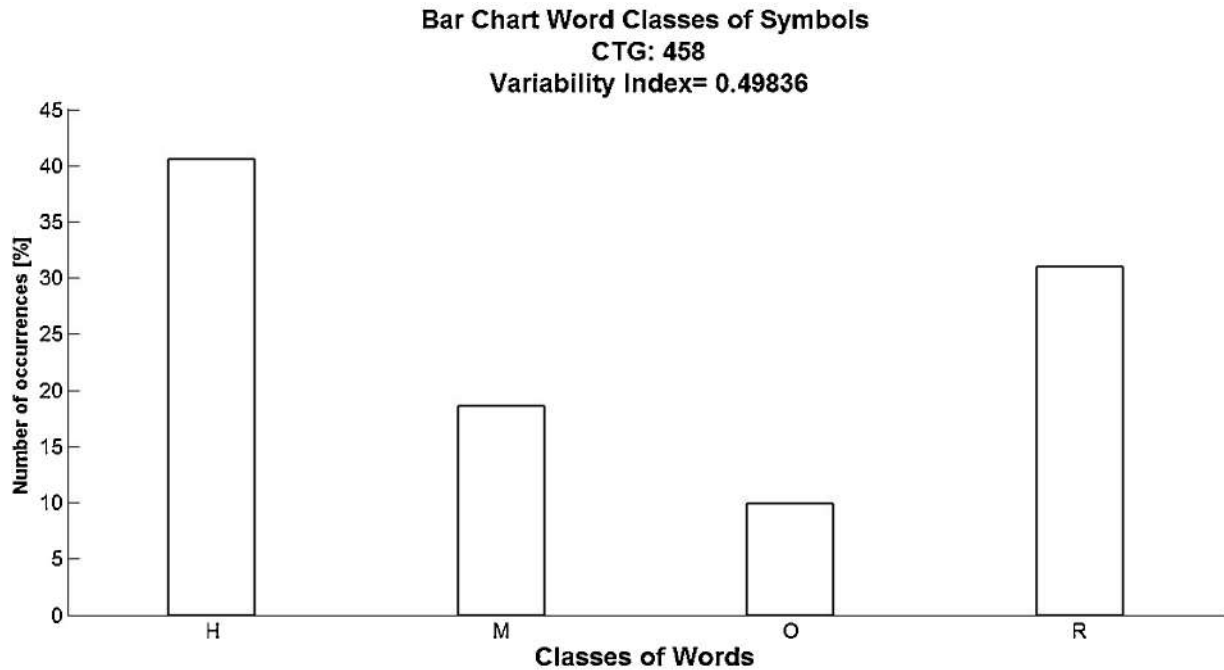


Fig. 10 – Distribution of word classes in FHR signal for CTG 458. V.I.=0.5.

which cannot be estimated with the traditional visual reading, but important to the diagnostic evaluation, like the approximate entropy [18]. The computerised CTG system described in their research work by Pardey et al. [63] realised a new method for the analysis of nonstress tests to provide a numerical definition of reactivity of the foetus that is tailored to the gestational age and independent on accelerations.

Despite to this wide availability, the interest in software developing is still current, also because there is the necessity for clinical or research purposes to develop an own system.

More recently, for example, Al-Yousif et al. [21] introduced a new software for digital CTG analysis to estimate the classical FHR parameters and, in 2012, Czabanski et al. [5] realised a predictor of neonatal acidemia risk by using a computerised software based on fuzzy analysis and classification of FHR signals to give a clinical interpretation of signals based on the FIGO guidelines.

In this perspective, we developed, for use in research, a software for antenatal CTG automatic analysis based on widely used criteria for visual analysis, which have the advantage of providing CTG interpretations familiar to obstetricians. Hence, in case, the software, due to this characteristic and to its simplicity and numerous functions, could be used also by clinicians for different aims, such as technical support or training.

Important characteristics of the software are that it is not developed specifically for a kind of cardiocardiographic device but it is compatible with the most part of commercial devices and that, in addition to the classical parameters, it involves indexes not conventional but potentially very useful in foetal monitoring as demonstrated also by preliminary results of our research group.

Besides, it is worth highlighting that this software has as first step a reliable pre-processing algorithm. This aspect is

somewhat neglected in the literature on the CTG; however an inadequate pre-processing of FHR signals may strongly affect analysis results, in particular the linear interpolation generates a consistent overestimation of SVB and artifacts removal is shown to be essential regardless of the spectral estimation technique because otherwise a decrease in PSD estimation and an error higher than 25% in STV assessment can be produced [14,28,42,64–66].

Another crucial but still controversial step in computerised CTG systems is the baseline estimation. Let us remember that the definition of many parameters depends on the baseline evaluation. Nevertheless, because of its qualitative definition (as well as for the floatingline), it is very difficult to verify whether or not the estimation of baseline is valid and, although specific works can be found in literature, the best technical procedure for baseline estimation is still controversial [19,67]. For these reasons, despite of limits inherent in baseline definition of Mantel, we decided to adopt his criteria, which are sufficiently detailed, defining a default procedure that we considered satisfactory but we leave to the obstetrician the possibility to interact with the system in order to modify the baseline smoothness.

Performances of our software (Tables 3 and 4) at first sight might seem unsatisfactory, but it is very important to point out that instead they are excellent when tested on simulated FHR signals (i.e. when the comparison with visual inspection was not necessary), confirming the essential differences in the procedure of events identification between a computerised system and human experts. Although we considered as gold standard for the comparison a team of very expert clinicians, we found an elevated inter-observer variability (Table 1). This inconsistency can be explained by different causes. Even if clinicians used FIGO guidelines, since the definitions are often ambiguous and still controversial, they are induced to

rely on his own experience. Another source of discrepancy between software and obstetricians analysis is the fact that clinicians often recognise only events with a high prognostic significance; in fact, even between them the disagreement is prevalent in the recognition of decelerations, whose meaning is very important and often debated also in literature [20,68,69]. To confirm these hypothesis, we found that, after a meeting aimed at defining in a more and more precise way the criteria for events inclusion, the agreement percentage for the parameter %50 was from 77.5 up to 85.8 for accelerations, from 50 up to 61.1 for decelerations and from 66.2 up to 71.4 for contractions. It is worth highlighting, finally, that clinicians, for their professional activity, are used making a global assessment of the signals, considering the set of all its features, so that the evaluation based on the detection of single events is more difficult; the performances of other software were in fact estimated on the basis of the overall evaluation of CTG [7,70].

In conclusion, we can evaluate satisfactory the performances preliminary of our software, which of course, will be subject of further improvement and extension and whose specifications will be update when a set of definitions really reproducible and standardised will be established. In any case, we planned the next version of the software by taking into account the following features: expanding the possibility of interaction between user and system, defining references range for indexes less classical, such as STV and V.I. (by SDA), introducing a post-processing, after events recognition, in order to include in the assessing the overall FHR variability and week of gestation.

Acknowledgements

This study was partially supported by the project “Sviluppo di sistemi per il monitoraggio fetale domiciliare” funded by “Legge 5” of Campania Region and by DRIVE IN2 project – funded by the Italian National Programme Piano Operativo Nazionale Ricerca e Competitività 2007/13.

The authors would like to acknowledge Dr. Mario Russo of the Gesan S.r.l. for the valuable support provided to the study.

They are very warmly thanking Francesca D’Alessandro and Alfonso Maria Ponsiglione for their gentle, continuous and passionate contribution to the development of software.

Finally, special recognition goes to the clinicians team (15 from “Dipartimento Ostetricia e Ginecologia, Cardarelli, Azienda Ospedaliera di Rilievo Nazionale” and 3 from “Dipartimento Materno Infantile, Ospedale Evangelico Villa Betania”, Naples, Italy) for the precious cooperation.

Appendix A. Technical details

In this appendix, we report technical details only of algorithms specifically developed and/or less common.

Before identification of all classical parameters, the FHR signal sections of bad quality are linearly interpolated.

All criteria for events detection, described in the following, were empirically established on the basis of literature indications and preliminary analysis of most part of CTG of our database.

A.1. Accelerations detection

The software identifies all FHR samples whose value is greater than the baseline (BSL) and calculates the maximum amplitude (measured in bpm) and the duration of these tracts (measured in seconds, s) exceeding the BSL. In addition, the area (measured in bpm × s) under the selected curve tract is preventively computed in order to avoid considering spikes and noisy tracts as possible accelerations. It is important to highlight that two consecutive tracts are considered together as one possible acceleration whether the space that separates them lasts less than 1 s and contains at least one sample greater than BSL + 15 bpm.

After that, the selected FHR tracts are considered as candidate accelerations whether they meet one of the following criteria:

- *First criterion for candidate acceleration:* if the tract contains at least a point greater than BSL + 15 bpm, the software identifies inside it all FHR samples greater than BSL + 5 bpm (not necessary consecutive) and computes the duration (measured in seconds and here called “d” for sake of brevity) of the tract between first and last identified sample.

If d is greater than 15 s and if the initial tract contains a sequence of consecutive FHR samples with value greater than BSL + 10 bpm that lasts more than 5 s, then the examined tract is considered a candidate acceleration.

- *Second criterion for candidate acceleration:* if the tract contains at least a point greater than BSL + 12 bpm, the software identifies inside it all FHR samples greater than the BSL (not necessary consecutive) and computes the corresponding duration (measured in seconds and called d).

If d is greater than 10 s and if the tract contains a sequence of consecutive FHR samples greater than BSL + 5 bpm that lasts more than 10 s, then the examined tract is considered a candidate acceleration.

- *Third criterion for candidate acceleration:* if the tract contains at least a point greater than BSL + 10 bpm, the software identifies inside it all FHR samples greater than the BSL (not necessary consecutive) and computes the duration (measured in seconds and called d) of the tract between first and last identified sample.

If d is greater than 20 s and if the tract contains a sequence of consecutive FHR samples with value greater than BSL + 10 bpm that lasts more than 10 s, then the examined tract is considered a candidate acceleration.

After computing all candidate accelerations, it is necessary to identify the presence of gaps within them.

A gap is defined as a tract of consecutive samples greater than the BSL but less than or equal to BSL + 5 bpm. In presence of a gap, two criteria are adopted, described below, for processing the candidate acceleration:

- *First criterion for gap:* if the gap lasts more than 10 s, it divides the candidate acceleration into separate parts. Each part remains accepted as acceleration only if its maximum amplitude is greater than BSL + 12 bpm and its duration is greater than 12 s at a level exceeding 5 bpm from the BSL. Alternatively, if all samples are greater than BSL + 12 bpm. Each part which does not meet these conditions is rejected as acceleration.

- *Second criterion for gap*: if the gap lasts less than 10 s, it divides the candidate acceleration into separate parts only if each separate part meets the requirements of the first criterion. Otherwise, the candidate acceleration is not divided. The last step for assessing accelerations consists in the control of interpolated values.
- *Criterion for interpolation*: if the candidate acceleration begins or ends with interpolated values, the starting or the ending point are respectively shifted to the first or the last sample of good quality. Then, the number of interpolated samples remaining inside the candidate acceleration is checked. If it does not percentage does not exceed 75% of acceleration duration, the acceleration is still maintained. Finally, if its samples are greater than $BSL + 10$ bpm for a duration of 15 s or more, the event under analysis is marked by the software as acceleration. As last step, the algorithm responsible for accelerations identification classifies them using other criteria, which, according to Mantel [19], include also the value of the area (expressed in $bpm \times s$).
- *Criterion for big accelerations*: are defined as big accelerations those with an area > 15 and a maximum amplitude > 20 bpm above BSL or area > 20 and amplitude > 15 .
- *Criterion for small accelerations*: small accelerations have an area > 12 and a maximum amplitude > 12 bpm above BSL or area > 15 and amplitude > 15 .
- *Criterion for very small accelerations*: are defined as very small accelerations, all accelerations that are not classified as big or small.

A.2. Decelerations detection

The software identifies all FHR samples whose value is less than the BSL and calculates the maximum amplitude (here meant as the minimum FHR value within the selected tract) and the duration of these tracts lying beneath the BSL. In addition, the area of the selected curve tract is preventively computed in order to avoid considering spikes and noisy tracts as possible decelerations. Two consecutive tracts are considered together as one possible deceleration whether the space that separates them lasts less than 2 s and contains at least one sample less than the BSL.

After that, the selected tracts are considered as candidate decelerations whether they meet one of the following criteria:

- *First criterion for candidate deceleration*: if the tract contains at least a point less than $BSL - 20$ bpm, the software identifies inside it all FHR samples less than the BSL (not necessary consecutive) and computes the duration (measured in seconds and here called d) of the tract between first and last identified sample. If d is greater than 30 s and if the initial tract contains a sequence of consecutive FHR samples with value less than $BSL - 10$ bpm that lasts more than 10 s, then the examined tract is considered a candidate deceleration.
- *Second criterion for candidate deceleration*: if the tract contains at least a point less than $BSL - 15$ bpm, the software identifies inside it all FHR samples less than the BSL (not necessary consecutive) and all FHR samples less than $BSL - 5$ (not necessary consecutive) and computes the corresponding

durations (measured in seconds and here called d_1 and d_2 respectively).

If d_1 is greater than 35 s and d_2 is greater than 25 s and if the tract contains a sequence of consecutive FHR samples less than $BSL - 5$ bpm that lasts more than 10 s, then the examined tract is considered a candidate deceleration.

- *Third criterion for candidate deceleration*: if the tract contains at least a point less than $BSL - 15$ bpm the software identifies inside it all FHR samples less than the BSL (not necessary consecutive) and all FHR samples less than $BSL - 5$ (not necessary consecutive) and computes the corresponding durations (measured in seconds and here called d_1 and d_2 respectively).

If d_1 is greater than 60 s and d_2 is greater than 30 s, then the examined tract is considered a candidate deceleration.

After computing all candidate decelerations, it is necessary to identify the presence of gaps within them.

In this case, a gap is defined as a tract of consecutive samples less than the BSL but greater than or equal to $BSL - 5$ bpm. In presence of a gap, two criteria are adopted, described below, for processing the candidate deceleration:

- *First criterion for gap*: if the gap lasts more than 10 s, it divides the candidate deceleration into two separate parts. Each part only remains accepted as deceleration if its maximum amplitude is less than $BSL - 12$ bpm and its duration is greater than 12 s at a level of 5 bpm below the BSL. Each part which does not meet these conditions is rejected as deceleration.
- *Second criterion for gap*: if the gap lasts less than 10 s, the candidate deceleration is not divided.

The last step for assessing decelerations consists in the control of interpolated values within the candidate deceleration.

- *Criterion for interpolation*: if the candidate deceleration begins or ends with interpolated values, the starting or the ending point are respectively shifted to the first or the last sample of good quality. Then, the number of interpolated samples remaining inside the candidate deceleration is checked. If its percentage does not exceed 70%, the deceleration is still maintained. Finally, if its samples are less than $BSL - 15$ bpm for a duration of 20 s or more, the event under analysis is marked by the software as deceleration.

As last step, the algorithm responsible for decelerations identification classifies them using other criteria, which, according to Mantel [19], include also the value of the area.

- *Criterion for big decelerations*: are defined as big decelerations those with an area > 20 and a maximum amplitude $< BSL - 20$ bpm.
- *Criterion for small decelerations*: small accelerations have an area > 15 , a maximum amplitude $< BSL - 15$ bpm.
- *Criterion for very small decelerations*: are defined as very small decelerations, all decelerations that are not classified as big or small.

A.3. Contractions detection

The software computes the basal tone (BT) of the UC signal (here called TOCO) by means of a procedure analogous to that used to evaluate the FHR baseline but using different filtering

Table 6 – Encoding to transform ΔRR into a symbols sequence.

ΔRR value	Symbol	Meaning
$\Delta RR > +ST$	P	High positive variation
$PT < \Delta RR < +ST$	p	Positive variation
$-PT < \Delta RR < +PT$	O	Absence of variability
$-ST < \Delta RR < -PT$	n	Negative variation
$\Delta RR < -ST$	N	High negative variation

parameters. After that, the software identifies all UC samples whose value is greater than the BT. Selected tracts are considered as candidate contractions whether they meet one of the following criteria:

- *First criterion for candidate contraction:* if the tract contains at least two points greater than $BT + 35$ mmHg, the software identifies inside it all TOCO samples greater than $BT + 5$ mmHg and computes the corresponding duration (measured in seconds and here called d1). Alternatively, the software identifies inside it all TOCO samples greater than $BT + 20$ mmHg and computes the corresponding duration (measured in seconds and here called d2). If d1 is greater than 30 s or d2 is greater than 10 s, then the examined tract is considered a contraction.
- *Second criterion for candidate contraction:* if the tract has a maximum amplitude greater than $BT + 25$ mmHg, the software identifies inside it all TOCO samples greater than the BT and all TOCO samples greater than $BT + 25$ mmHg and computes the corresponding durations (measured in seconds and here called d1 and d2 respectively). If d1 is greater than 45 s and d2 is greater than 6 s, then the examined tract is considered a contraction. As last step, the algorithm responsible for contractions identification classifies them using other criteria:
- *Criterion for big contractions:* are defined as big contractions those with an amplitude >35 mmHg above BT for at least 45 s.
- *Criterion for small contractions:* are defined as small contractions, all contractions that are not classified as big.

A.4. Symbolic dynamics analysis

Starting from FHR signals (in bpm), the software calculates the RR series (in ms) using the formula: $RR = 60,000/FHR$. From the RR time series, the series of the differences between consecutive heart beat intervals (ΔRR) is computed, to make the analysis independent by the mean heart rate. Then, a transformation into symbols is carried on. To this aim, an alphabet of five symbols, defined by the authors, is used to classify the full range of variability of the ΔRR series. In order to distinguish higher from moderate variations, the decision rule that associates each series sample with one alphabet symbol is the comparison between the sample value and the values of two thresholds, named primary (PT) and secondary (ST) threshold (Table 6). The primary threshold (PT) identifies samples that are not due to the physiological variability of the cardiac rhythm but to the noise (mainly correlated to the resolution in FHR recording). ST is heuristically set to 3 ms, by considering the peak to peak amplitude of FHRV. Then, a sliding window of length L is shifted along the codified series, with

Table 7 – Criterion adopted to group the words in WC.

Criterion	Code	Meaning
At least 4 symbols “P” or “N”	H	High activation
At least 3 symbols “P” and 1 symbol “p”		
At least 3 symbols “N” and 1 symbol “n”	M	Moderate activation
At least 4 symbols “p” or “n”		
At least 3 symbols “p” and 1 symbol “P”		
At least 3 symbols “n” and 1 symbol “N”	O	Absence of variability
At least 4 symbols “O”		
All other cases	R	Random

an overlap of $L - 1$ samples, transforming it into a sequence of patterns of L samples (called words). The L value is equal to 7 (considering a mean foetal heart rate of 140 bpm, this value corresponds to 3 s) in order to surely include in a one word the burst peak of a sympathetic response. At this point, a classification scheme is necessary to describe the obtained words. According to a method previously introduced [41–43], the words were grouped in different word classes (WC) by the within-word symbol occurrence, using the criterion (we called also dominance criterion) described in Table 7.

The criterion has the aim to highlight sympatho-vagal activations, considered associated to greater or less variability, and cases of absence of variability or random variations. Then, the frequency distributions of WC is analysed. Finally, a novel variability index (V.I.) is estimated from percentages of occurrence of the different WC (pH, pM, pO and pR), with the aim to put in evidence the amount of physiological variability of the signal at the expense of that null or random, at which we assign zero weight:

$$V.I. = pH + pM \times 0.5 + (pO + pR) \times 0 \quad (4)$$

As last step, in order to quantify FHRV, three ranges of values were experimentally set for V.I.: $V.I. < 0.20$ – low variability, $0.20 \leq V.I. \leq 0.28$ – medium variability, $V.I. > 0.28$ – high variability.

REFERENCES

- [1] R.M. Grivell, Z. Alfrevic, G.M. Gyte, D. Devane, Antenatal cardiotocography for fetal assessment, *Cochrane Database Syst. Rev.* 1 (2010).
- [2] R. Goeshen, *Cardiotocografia pratica*, V edizione CIC Edizioni Internazionali, Roma, 1998.
- [3] J.T. Curran, *Fetal Heart Monitoring*, Butterworths, 1975.
- [4] T. Tongsong, A. Iamthongin, C. Wanapirak, W. Piyamongkol, S. Sirichotiyakul, P. Boonyanurak, C. Neelasri, Accuracy of fetal heart-rate variability interpretation by obstetricians using the criteria of the National Institute of Child Health and Human Development compared with computer-aided interpretation, *J. Obstet. Gynaecol. Res.* 31 (1) (2005) 68–71.
- [5] R. Czabanski, J. Jezewski, A. Matonia, M. Jezewski, Computerized analysis of fetal heart rate signals as the predictor of neonatal acidemia, *Expert Syst. Appl.* 39 (15) (2012) 11846–11860.

- [6] A. Sweha, T.W. Hacker, J. Nuovo, Interpretation of the electronic fetal heart rate during labor, *Am. Fam. Physician* 59 (1999) 2487–2506.
- [7] D. Ayres-de-Campos, J. Bernardes, A. Garrido, J. Marques-de-Sa, L. Pereira-Leite, SisPorto 2.0: a program for automated analysis of cardiotocograms, *J. Matern. Fetal Neonatal Med.* 9 (5) (2000) 311–318.
- [8] R. Mantel, H.P. Van Geijn, I.A. Ververs, G.J. Colenbrander, P.J. Kostense, Automated analysis of antepartum fetal heart rate in relation to fetal rest-activity states: a longitudinal study of uncomplicated pregnancies using the Sonicaid System 8000, *Eur. J. Obstetr. Gynecol. Reprod. Biol.* 71 (1) (1997) 41–51.
- [9] D.K. Donker, H.P. van Geijn, A. Hasman, Interobserver variation in the assessment of fetal heart rate recordings, *Eur. J. Obstetr. Gynecol. Reprod. Biol.* 52 (1993) 21–28.
- [10] J. Bernardes, A. Costa-Pereira, D. Ayres-de-Campos, H.P. van Geijn, L. Pereira-Leite, Evaluation of interobserver agreement of cardiotocograms, *Int. J. Gynecol. Obstetr.* 57 (1997) 33–37.
- [11] J.B. Trimbois, M.J.N.C. Keirse, Observer variability in assessment of antepartum cardiotocograms, *BJOG: Int. J. Obstetr. Gynaecol.* 85 (12) (1978) 900–906.
- [12] T. Todros, C.U. Preve, C. Plazzotta, M. Biolcati, P. Lombardo, Fetal heart rate tracings: observers versus computer assessment, *Eur. J. Obstetr. Gynecol. Reprod. Biol.* 68 (1996) 83–86.
- [13] Royal College Of Obstetricians and Gynaecologists (RCOG), (2001). The use of electronic fetal monitoring. The use and interpretation of cardiotocography in intrapartum fetal surveillance. R. Coll. Obstet. Gynaecol. Evidence-based Clinical Guideline Number 8.
- [14] M. Cesarelli, M. Romano, P. Bifulco, F. Fedele, M. Bracale, An algorithm for the recovery of fetal heart rate series from CTG data, *Comput. Biol. Med.* 37 (5) (2007) 663–669.
- [15] T.P. Sartwelle, Electronic fetal monitoring: a bridge too far, *Leg. Med.* 33 (2012) 313–379 (Parsons, L.L.P., Houston, TX 77056, USA).
- [16] D. Ayres-de-Campos, C. Costa-Santos, J. Bernardes, Prediction of neonatal state by computer analysis of fetal heart rate tracings: the antepartum arm of the SisPorto multicentre validation study, *Eur. J. Obstetr. Gynecol. Reprod. Biol.* 118 (1) (2005) 52–60.
- [17] B.K. Strachan, D.S. Sahota, W.J. van Wijngaarden, D.K. James, M.Z. Chang, Computerised analysis of the fetal heart rate and relation to acidaemia at delivery, *Br. J. Obstet. Gynaecol.* 108 (August) (2001) 848–852.
- [18] G. Magenes, M.G. Signorini, M. Ferrario, F. Lunghi, 2CTG2: A new system for the antepartum analysis of fetal heart rate, in: 11th Mediterranean Conference on Medical and Biomedical Engineering and Computing 2007, Springer Berlin Heidelberg, 2007, pp. 781–784.
- [19] D. Ayres-de-Campos, J. Bernardes, Comparison of fetal heart rate baseline estimation by SisPorto 2.01 and a consensus of clinicians, *Eur. J. Obstetr. Gynecol. Reprod. Biol.* 117 (2) (2004) 174–178.
- [20] J. Jezewski, J. Wrobel, K. Horoba, A. Matonia, T. Kupka, Estimation of beat-to-beat accuracy of fetal heart rate data obtained via Doppler ultrasound, in: Proceedings of Embec'02 2nd European Medical and Biological Engineering Conference, December 04–08, 2002, Vienna, Austria, 2002, II/1536.
- [21] S.N. Al-Yousif, M.M. Ali, Cardiotocography trace pattern evaluation using MATLAB Program, in: Proceedings of International Conference on Biomedical Engineering and Technology (ICBET 2011), 2011.
- [22] P. Van Leeuwen, D. Cysarz, S. Lange, D. Geue, D. Groenemeyer, Quantification of fetal heart rate regularity using symbolic dynamics, *Chaos: Interdiscip. J. Nonlinear Sci.* 17 (1) (2007) 015119.
- [23] D. Cysarz, P. Van Leeuwen, H. Bettermann, Irregularities and nonlinearities in fetal heart period time series in the course of pregnancy, *Herzschrittmachertherapie und Elektrophysiologie* 11 (3) (2000).
- [24] A. Voss, S. Schulz, R. Schroeder, M. Baumert, P. Caminal, Methods derived from nonlinear dynamics for analysing heart rate variability, *Phil. Trans. R. Soc. A* 367 (2009) 277–296.
- [25] S. Guzzetti, E. Borroni, P.E. Garbelli, E. Ceriani, P. Della Bella, N. Montano, C. Cogliati, V.K. Somers, A. Mallani, A. Porta, Symbolic dynamics of heart rate variability. A probe to investigate cardiac autonomic modulation, *Circulation* (July) (2005).
- [26] R. Mantel, H.P. Van Geijn, F.J.M. Caron, J.M. Swartjes, E.E. Van Woerden, H.W. Jongswa, Computer analysis of antepartum fetal heart rate: 1. Baseline determination, *Int. J. Biomed. Comput.* 25 (4) (1990) 261–272.
- [27] R. Mantel, H.P. Van Geijn, F.J.M. Caron, J.M. Swartjes, E.E. Van Woerden, H.W. Jongswa, Computer analysis of antepartum fetal heart rate: 2. Detection of accelerations and decelerations, *Int. J. Biomed. Comput.* 25 (4) (1990) 273–286.
- [28] M. Romano, G. Faiella, P. Bifulco, G. D'Addio, F. Clemente, M. Cesarelli, Outliers detection and processing in CTG monitoring, in: XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013, Springer International Publishing, 2014, pp. 651–654.
- [29] G. Improta, M. Romano, F. Amato, M. Sansone, M. Cesarelli, Development of a software for automatic analysis of CTG recordings. GNB2012, June 26th–29th, 2012.
- [30] FIGO NEWS, Guidelines for the use of fetal monitoring, *Gynaecol. Obstetr.* (1986).
- [31] NICHD. Research Planning Workshop, Electronic fetal heart rate monitoring: research guidelines for interpretation. National Institute of Child Health and Human Development Research Planning Workshop, *Am. J. Obstet. Gynecol.* 177 (1997) 1385–1390.
- [32] D. Arduini, G. Rizzo, F. Giannini, G.G. Garzetti, C. Romanini, Computerized analysis of fetal heart-rate. 2. Comparison with the interpretation of experts, *J. Maternal-Fetal Investig.* 3 (3) (1993) 165–168.
- [33] S.P. Von Steinburg, A.L. Boulesteix, C. Lederer, S. Grunow, S. Schiermeier, W. Hatzmann, M. Daumer, What is the normal fetal heart rate? *Peer J.* 1 (2013) e82.
- [34] C.W.G. Redman, SonicaidFetalCare. Clinical Application Guide, Oxford Instruments Medical Ltd, 2003.
- [35] M. Cesarelli, M. Romano, G. D'Addio, M. Ruffo, P. Bifulco, G. Pasquariello, A. Fratini, Floatingline estimation in FHR signal analysis, in: 5th European Conference of the International Federation for Medical and Biological Engineering, Springer Berlin Heidelberg, 2012, pp. 179–182.
- [36] M. Cesarelli, M. Romano, P. Bifulco, Comparison of short term variability indexes in cardiotocographic foetal monitoring, *Comput. Biol. Med.* 39 (2) (2009) 106–118.
- [37] M. Romano, P. Bifulco, M. Cesarelli, M. Sansone, M. Bracale, Foetal heart rate power spectrum response to uterine contraction, *Med. Biol. Eng. Comput.* 44 (3) (2006) 188–201.
- [38] M. Cesarelli, M. Romano, M. Ruffo, P. Bifulco, G. Pasquariello, Foetal heart rate variability frequency characteristics with respect to uterine contractions, *Commun. Netw.* 2 (4) (2010).
- [39] M. Romano, M. Bracale, M. Cesarelli, M. Campanile, P. Bifulco, M. De Falco, A. Di Lieto, Antepartum cardiotocography: a study of fetal reactivity in frequency domain, *Comput. Biol. Med.* 36 (6) (2006) 619–633.
- [40] J. Karin, M. Hirsch, C. Sagiv, S. Akselrod, Fetal autonomic nervous system activity monitoring by spectral analysis of heart rate variations, in: Proceedings of Computers in Cardiology 1992, IEEE, 1992, pp. 479–482.

- [41] M. Romano, M. Cesarelli, P. Bifulco, M. Ruffo, A. Fratini, G. Pasquariello, Time-frequency analysis of CTG signals, *Curr. Dev. Theory Appl. Wavelets* 3 (2) (2009) 169–192.
- [42] M. Cesarelli, M. Romano, M. Ruffo, P. Bifulco, G. Pasquariello, A. Fratini, PSD modifications of FHRV due to interpolation and CTG storage rate, *Biomed. Signal Process. Control* 6 (3) (2011) 225–230.
- [43] L.W. Oppenheimer, R.M. Lewinsky, 7 Power spectral analysis of fetal heart rate, *Baillière's Clin. Obstetr. Gynaecol.* 8 (3) (1994) 643–661.
- [44] S. Cerutti, S. Civardi, A. Bianchi, M.G. Signorini, E. Ferrazzi, G. Pardi, Spectral analysis of antepartum heart rate variability, *Clin. Phys. Physiol. Meas.* 10 (4B) (1989) 27.
- [45] L.J. Groome, D.M. Mooney, L.S. Bentz, K.P. Singh, Spectral analysis of heart rate variability during quiet sleep in normal human fetuses between 36 and 40 weeks of gestation, *Early Hum. Dev.* 38 (1) (1994) 1–9.
- [46] 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 (3) (2003) 365–374.
- [47] P. Laguna, G.B. Moody, R.G. Mark, Power spectral density of unevenly sampled data by least-square analysis: performance and application to heart rate signals, *IEEE Trans. Biomed. Eng.* 45 (6) (1998) 698–715.
- [48] N.R. Lomb, Least-squares frequency analysis of unequally spaced data, *Astrophys. Space Sci.* 39 (2) (1976) 447–462.
- [49] J.D. Scargle, Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data, *Astrophys. J.* 263 (1982) 835–853.
- [50] B.P. Flannery, W.H. Press, S.A. Teukolsky, W. Vetterling, *Numerical Recipes in C*, Press Syndicate of the University of Cambridge, New York, 1992.
- [51] A. Porta, G. D'Addio, G.D. Pinna, R. Maestri, T. Gnechi-Ruscione, R. Furlan, . . . A. Malliani, Symbolic analysis of 24 h Holter heart period variability series: comparison between normal and heart failure patients, *Comput. Cardiol.* (September 2005) 575–578 (IEEE).
- [52] A. Voss, J. Kurths, H.J. Kleiner, A. Witt, N. Wessel, P. Saparin, R. Dietz, The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death, *Cardiovasc. Res.* 31 (3) (1996) 419–433.
- [53] M. Romano, P. Bifulco, G. Improta, G. Faiella, M. Cesarelli, F. Clemente, G. D'Addio, Symbolic dynamics in cardiotocographic monitoring, in: *E-Health and Bioengineering Conference (EHB)*, November 2013, IEEE, 2013, pp. 1–4.
- [54] M. Cesarelli, M. Romano, P. Bifulco, G. Improta, G. D'Addio, Prognostic decision support using symbolic dynamics in CTG monitoring, *Stud. Health Technol. Inform.* 186 (2013) 140–144.
- [55] M. Cesarelli, M. Romano, P. Bifulco, G. Improta, G. D'Addio, An application of symbolic dynamics for FHRV assessment, *Stud. Health Technol. Inform.* 180 (2012) 123–127.
- [56] G. Improta, M. Romano, A.M. Ponsiglione, P. Bifulco, G. Faiella, M. Cesarelli, Computerized cardiotocography: a software to generate synthetic signals, *J. Health Med. Inform.* 5 (4) (2014), ISSN: 2157-7420.
- [57] J.T. Parer, E.F. Hamilton, Comparison of 5 experts and computer analysis in rule-based fetal heart rate interpretation, *Am. J. Obstet. Gynecol.* 203 (5) (2010) 451–461.
- [58] M. Mongelli, R. Dawkins, T. Chung, D. Sahota, J.A. Spencer, A.M. Chang, Computerised estimation of the baseline fetal heart rate in labour: the low frequency line, *BJOG: Int. J. Obstetr. Gynaecol.* 104 (10) (1997) 1128–1133.
- [59] J. Kottner, L. Audigé, S. Brorson, A. Donner, B.J. Gajewski, A. Hrobjartsson, C. Roberts, M. Shoukri, D.L. Streiner, Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed, *J. Clin. Epidemiol.* 64 (2011) 96–106.
- [60] T. Zhu, A.E.W. Johnson, J. Behar, G.D. Clifford, Bayesian voting of multiple annotators for improved QT interval estimation, in: *Computing in Cardiology Conference (CinC)*, 2013, pp. 659–662, FE.
- [61] J. Špilka, V. Chudáček, P. Janku, L. Hruban, M. Burša, M. Huptych, L. Zach, L. Lhotská, Analysis of obstetricians' decision making on CTG recordings, *J. Biomed. Inform.* (2014), <http://dx.doi.org/10.1016/j.jbi.2014.04.010> (Published online: 15.04.14).
- [62] 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. Perinatal Med.* 36 (3) (2008) 260–264.
- [63] J. Pardey, M. Mouldenb, C.W. Redman, A computer system for the numerical analysis of nonstress tests, *Am. J. Obstet. Gynecol.* 186 (5) (2002) 1095–1103.
- [64] M.P. Tarvainen, J.P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, P.A. Karjalainen, Kubios HRV – heart rate variability analysis software, *Comput. Methods Prog. Biomed.* 113 (1) (2014) 210–220.
- [65] G.D. Clifford, L. Tarassenko, Quantifying errors in spectral estimates of HRV due to beat replacement and resampling, *IEEE Trans. Biomed. Eng.* 52 (April (4)) (2005).
- [66] H. Goncalves, A.P. Rocha, D. Ayres-De-Campos, J. Bernardes, Internal versus external intrapartum foetal heart rate monitoring: the effect on linear and nonlinear parameters, *Physiol. Meas.* 27 (2006) 307–319.
- [67] A. Georgieva, S.J. Payne, M. Moulden, C.W. Redman, Computerized fetal heart rate analysis in labor: detection of intervals with un-assignable baseline, *Physiol. Meas.* 32 (10) (2011) 1549.
- [68] M. Kazandi, F. Sendag, F. Akercan, M.C. Terek, G. Gundem, Different types of variable decelerations and their effects to neonatal outcome, *Singapore Med. J.* 44 (5) (2003) 243–247.
- [69] J. Chuang, C.T. Chou, W.C. Cheng, L.W. Huang, J.L. Hwang, Y.L. Tsai, Spontaneous fetal heart rate deceleration: an ominous sign for fetal outcome, *Arch. Gynecol. Obstet.* 269 (4) (2004) 254–258.
- [70] S. Schiermeier, A. Pildner von Steinburg, J. Thieme, M. Reinhard, M. Daumer, W. Scholz, K.T.M. Hatzmann, Schneider, Sensitivity and specificity of intrapartum computerised FIGO criteria for cardiotocography and fetal scalp pH during labour: multicentre, observational study. RCOG 2008, *BJOG: Int. J. Obstetr. Gynaecol.* (2008) 1557–1563.