SUITABILITY OF FRACTAL DIMENSION ANALYSIS OF FOETAL HEART RATE AS AN INDICATOR FOR ASPHYXIA

P A Hopkins
University of Plymouth, U.K.

N J Outram
University of Plymouth, U.K.

E C Ifeachor
University of Plymouth, U.K.

R Sneyd
Peninsula Medical School, U.K

K G Rosén
University of Gothenburg and
Neoventor Medical AB, Sweden

ABSTRACT

There are potential healthcare benefits to be gained from the real time monitoring of Heart Rate during critical periods such as labour, surgery or neonatal intensive care. Fractal dimension is a measure of system complexity. This paper assesses the ability of three methods for estimation of fractal dimension to distinguish between healthy and compromised FHR data. The techniques used are the box counting dimension, Katz's Algorithm and Sevcik's Algorithm.
It is found that all three techniques can distinguish between healthy and compromised FHR data, however these distinctions are not always apparent when comparing short blocks of data. In order to improve specificity it is necessary to use longer time blocks, which may not be available continuously within a dataset.

This work was supported by the EPSRC and the EU BIOPATTERN Network of Excellence.

Keywords: Fetal monitoring, ECG, heart rate, FECG, complexity, fractal dimension

I. INTRODUCTION

A consolidation of all available phenotype and genotype data for a given individual is what we term a Bioprofile [1]. An "EARLYLIFE Bioprofile" would be one that belongs to a child or foetus. Appropriate access to such a resource would ultimately enable individualised healthcare. Researchers with access to population based Bioprofile data could discover new knowledge about the causes, prevention and treatment of diseases and conditions. Our long term work within the BIOPATTERN project [1] aims is to address many of the issues specific to building an EARLYLIFE Bioprofile and to enable the discovery and validation of new knowledge that could enable individualised healthcare in the future.

Adverse events in early life, whether during gestation, childbirth or as a neonate, can have a profound impact on long term neurological development and/or health of the individual [2]. In this paper we are focusing on detection and characterisation of adverse events during labour, such as birth asphyxia, which can result in brain injury [3] and/or medical intervention. Apart from the immediate problems related to neonatal morbidity and mortality, lack of oxygenation increases the individuals' risk of developmental disorders such as cerebral palsy and possibly learning difficulties, behavioural problems and other diseases and conditions. In some cases these can go undetected for many years thereby missing the opportunity for early treatment.

Electronic Foetal Monitoring (EFM) is used to assess the condition of the foetus during labour and to prevent situations that lead to neurological damage. Standard EFM is based on a continuous presentation of the heart-rate and uterine contractions, known as the CTG. Unfortunately this has not been proven to improve the outcome of labour, and is often attributed to contributing to a rise in intervention rates retrospectively deemed unnecessary. To address this, a new technology, the STAN® monitor, provides additional information about ST waveform changes in the foetal ECG. Alongside the CTG, this has been shown to safely reduce the incidence of birth related asphyxia as well as the intervention rate [4]. However, this still requires human analysis of the HR patterns in the CTG, and there are still some rare cases whereby the abnormal CTG patterns are visually very subtle and are sometimes missed. Clinically, the most important pattern is reduced beat-to-beat variation occurring over time. This paper will focus on techniques for more objective assessment including detection and classification of changes in FHR variability (FHRV) recorded during labour in a non-stationary FHR signal. The non-stationary nature of the FHR signal during uterine contractions in particular has so far hindered continuous assessment of FHRV [15]. If successful, this could be used to prevent further cases of brain injury and hence developmental problems in later life. Such a system would provide valuable Biopattern data which could be added to each individual's Bioprofile and may provide prognostic as well as diagnostic information to help tailor healthcare to an individuals needs.

The heart rate is regulated by the cardio vascular system (CVS) and autonomic nervous system (ANS). Evidence of rapid beat-to-beat modulation of the heart rate, known as heart rate variability (HRV), is indicative of a foetus with an ANS that is appropriately reacting to the stress and dynamics of labour [5]. A reduction in HRV, in the absence of other indicators (such as heart rate accelerations), is an indication of abnormality as it implies a compromise in ANS function. It is difficult however to assess HRV by visual interpretation. To complicate matters further, a sleep state or the effects of drugs administered during labour will result in a short term reduction in 'visual' HRV. A continuous quantitative metric is clearly required to enable more sensitive and specific diagnosis.

Fractal Dimension (FD) is a measure for describing a systems complexity from observations of its output. The concept of dimension applied to HR is also attractive as it goes some way to describe the state of the underlying dynamical systems that regulate the heart, i.e. the ANS. If successful, this could potentially be related back to a more physiological description of the foetal state than the somewhat abstract CTG does today. It is also interesting to note that it has been suggested that normal FHR is a chaotic system with fractal nature [6].

We expect a reactive healthy foetus to have a complex HR trace. Complex systems will have higher dimension than 'simple' systems. Thus, FD of fetal HR in a healthy individual would be expected to be higher than that for a compromised individual with a depressed ANS function. This has been demonstrated in both adults [7] and neonates [8]. FD may be better suited to FHR monitoring/analysis than other non-linear techniques such as correlation dimension and Lyapunov exponents as it can be calculated from smaller datasets allowing greater resolution in time, thus it handles the non-stationary nature of the signal. Furthermore, FD estimates are robust with respect to missing data which is a common problem in EFM.

The classic Hausdorff-Besicovitch definition of dimension, given in equation 1, assumes a function that can be analysed at ever decreasing scale (i.e. it is possible to zoom into any region of an object / function to examine detailed structure). This is clearly possible for continuous functions, but not generally (or at least limited) for sampled signals. It is even less appropriate for HR signals where the system (ANS and CVS) essentially produce a sampled signal that is irregular and coarse. We have therefore compared a number of more practical *estimates* of dimension for the analysis of abnormal and 'challenging' normal cases. These return values of dimension which approximate to the Hausdorff Besicovitch dimensional value. The objective is to evaluate the suitability of different estimates for clinical value rather than faithful estimation of dimension.

The paper is organised as follows. Section II introduces FD theory. An overview of the FD algorithms investigated in this paper is presented in Section III. In Section IV, the method used for this study is presented. Results and discussions are presented in Section V. Concluding remarks and future directions in this research are given in Section VI.

II. Fractal Dimension

Fractal dimension is a measure of system complexity**.** Familiar geometric objects such as a line or a square have an integer dimension value in Cartesian space e.g. a line is one dimensional, a square is two dimensional etc. A definition of dimension that extends this concept whilst remaining compatible with traditional geometry is that of the Hausdorff-Besicovitch dimension [9]. Chaotic and highly non-linear systems are said to have *non-integer* dimension greater than the topological dimension. An example is a fractal which is formed by recursive iterations of simple functions with feedback. These have the interesting quality of being self-similar on many scales e.g. if a small part of the set is examined its form will resemble that of the entire set.

An estimate of fractal dimension can determine if a system is chaotic from observed outputs (such as a time series). There are many techniques used to estimate FD but fundamentally what is determined is how its area (for a shape), or length (in the case of a line), changes as it is measured at different scales. For example any line of length L can be completely covered by non-overlapping circles with radius ε. $N(\varepsilon)$ is the number of such circles necessary to cover the set. Thus $N(\varepsilon)$ is a measure of the length of the line at scale ε. If ε is reduced in steps and the length of the line is measured for each ε the dimension of the line will be given by

$$D_h = \lim_{\varepsilon \to 0} - \frac{\ln N(\varepsilon)}{\ln(\varepsilon)} \tag{1}$$

This is a basic definition of the Hausdorff-Besicovitch dimension D_h. 'A fractal is by definition a set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological dimension' [10]. This definition of the Hausdorff-Besicovitch dimension can be expanded to cover shapes and volumes of higher dimension by covering with spheres of higher dimension.

For a straight or simple line, as ε gets smaller, in the limit $N(\varepsilon)=1/\varepsilon$ and $D_h=1$. For a fractal however, such as Koch's curve, the length $N(\varepsilon)$ increases as a constant power law relationship with each scale decrease

$$N(\varepsilon) = Ke^{-D_h} \tag{2}$$

where $D_h>1$. A true fractal will hold this constant power law relationship over infinite decreases in scale. However for sampled real world data such as FHR we do not have access to continuous functions thus measuring the limit as $\varepsilon \to 0$ is impossible. Thus while the Hausdorff-Besicovitch dimension is robust and consistent with the topological dimension of regular geometric shapes, it is not suitable for practical application. For this reason there exist several different techniques for the estimation of dimension which although not as precise, can be applied to real world data.

Many estimators of fractal dimension have been developed in order to classify discrete time series data. Most of this data exists in 'affine space' where there is no relationship between the axes of a plot e.g. voltage versus time. This can create problems with maintaining consistency. However, FHR data is a time series derived from the time-intervals between successive beats of a heart, known as RR intervals, such that HR=60,000/RR, where RR is measured in ms. We therefore choose to work with RR intervals as the units are the same on each axis.

III. FRACTAL DIMENSION ALGORITHMS

Box Counting Dimension

Assuming a set lies in an N dimensional Cartesian space. This space can then be divided into a grid of N dimensional sub spaces with edge length ε. For the purpose of this work, the 'set' is the set of all the RR interval measurements. The plot of RR intervals against time lies on a plane (paper surface), thus $N=2$ and the sub spaces are simply squares with side length ε. The number of subspaces $\tilde{N}(\varepsilon)$ necessary to cover the set is then counted. This process is repeated for successively smaller values of ε. The box counting dimension d should link $\tilde{N}(\varepsilon)$ and ε such that [9].

$$\tilde{N}(\varepsilon) \sim 1/\varepsilon^{d} \tag{3}$$

This is only the case if there exists a constant k such that $\lim_{\varepsilon\to 0} \tilde{N}(\varepsilon) / \left(1/\varepsilon^{d} \right) = k$. Solving for d gives

$$d = \lim_{\varepsilon\to 0} \left(\frac{\ln k - \ln \tilde{N}(\varepsilon)}{\ln \varepsilon} \right) \tag{4}$$

In the limit as $\varepsilon \to 0$, this can be simplified to

$$d = -\lim_{\varepsilon\to 0} \left(\frac{\ln \tilde{N}(\varepsilon)}{\ln \varepsilon} \right) \tag{5}$$

Note the similarity to the Hausdorff-Besicovitch definition in equation 1. The algorithm used is developed from Boshoff's fast box counting algorithm [11].

Katz's Algorithm

Katz proposed that an estimate for the fractal dimension of a set can be derived from following equation [12].

$$D_k = \frac{\log_{10}(L)}{\log_{10}(d)} \tag{6}$$

where d is the maximum distance from the first to the furthest point in the set and L is the total length of the set, obtained by summing the distances between i successive points. In order to take into account the possibility of differing scale on the axes a normalisation factor is introduced. Equation 6 can be written

$$D_k = \frac{\log_{10}(L/a)}{\log_{10}(d/a)} \tag{7}$$

where a is the average number of steps in the series. If n is the number of steps in the curve then $n=L/a$ and

$$D_k = \frac{\log_{10}(n)}{\log_{10}(d/L) + \log_{10}(n)} \tag{8}$$

For FHRMD data computational complexity can be reduced by assuming that d is equal to the data span on the abscissa since the variation in the ordinate is relatively very small for a time series of any length.

Sevcik's Algorithm

This equation is developed from the expression for the Hausdorff –Besicovitch dimension (see equation 1). For any value of ε, $N(\varepsilon) = L/2\varepsilon$. Thus the equation for D_h can be written [13]

$$\begin{aligned} D_h &= \lim_{\varepsilon\to 0} \left[\frac{-\ln(L) + \ln(2\varepsilon)}{\ln(\varepsilon)} \right] \\ &= \lim_{\varepsilon\to 0} \left[1 - \frac{\ln(L) - \ln(2)}{\ln(\varepsilon)} \right] \\ &= \lim_{\varepsilon\to 0} \left[1 - \frac{\ln(L)}{\ln(\varepsilon)} \right] \end{aligned} \tag{9}$$

A time series of N data values, $x_i : y_i$, sampled between x =0 and x=$t_{\max}$ can be mapped on to a unit square by applying the following linear transforms:

$$x_i^{*} = \frac{x_i}{x_{\max}} \tag{10}$$

$$y_i^{*} = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \tag{11}$$

where x_i^{*} and y_i^{*} are the normalised abscissa and ordinate values respectively, x_{max} is the maximum x_i value and y_{min} and y_{max} are the maximum and minimum y_i values. For N values of a time series mapped to the unit square, by letting $\varepsilon=1/(2N')$ the fractal dimension (Φ) can then be approximated by D where

$$\Phi \approx D = 1 + \frac{\ln(L)}{\ln(2N')} \tag{12}$$

where $N' = N$-1, and L is the Euclidean length of the curve in the unit square. The purpose of mapping to a unit square is to overcome the problem with axes in affine space. Since FHRMD data has identical time axes it is not necessary to apply the mapping function to data before processing.

Modified Sevcik Algorithm

It can be seen from equation 13 that the mapping function for the abscissa relies on the value of y_{min} and y_{max}. For real time monitoring these values are not known a priori. However, the normal range of HR is known (60-240bpm). An offset of +250ms was applied to each data value, and the range 0-500ms was then mapped to the ordinate of the unit square such that.

$$y_i^* = \frac{y_i - 250ms}{500ms} \tag{13}$$

IV. METHOD

Sample FHR data was selected from a clinical database of some 20,000 cases according to medical outcome. This consisted of thirty 'normal' cases with normal fetal scalp pH obtained during 1st stage of labour and normal cord artery blood gases and Apgar scores at delivery. Four index cases with adverse outcome due to intrapartum asphyxia were also selected (others have been retained for validation purposes). Data was obtained using the STAN®S 21 system as part of a multicentre project on intrapartum foetal monitoring [16]. The data used were the RR intervals derived from foetal ECG. The STAN system only records RR intervals when the raw signal is of sufficient quality. Inevitably a number of data samples are dropped so this needs to be handled safely. The FD is based on analysis of contiguous 2 minute blocks. These are only processed if the total duration of missing data does not exceed 500ms.

The foetal heart rate is considered to be the superposition of a primary heart component (the slow trend) and the FHR modulating determinants (FHRMD) which are the short-term beat-to-beat changes in the FHR pattern attributed to changes in the ANS [14]. It is the FHRMD we are interested in, so this is separated as follows:
(i) Data in each two minute block is first linearly interpolated and evenly sampled every 10 milliseconds; (ii) the primary FHR (PFHR) component is approximated by fitting a piecewise polynomial function (with continuity of up to second derivative) to the interpolated RR interval data; (iii) the PHFR values are then subtracted from the interpolated RR data samples to obtain a *residual* vector which represents the FHRMD.
For each case, FD is estimated for each 2 minute block of residual data and recorded in a table. This was then repeated for each different method of FD estimation. In addition, simple standard deviation of 2-minute blocks was added as an additional metric. The preliminary results are now discussed.

V. RESULTS

Fractal Dimension values for two minute blocks were plotted for each case. From these plots it was apparent that the index (abnormal outcome) cases consistently exhibit a lower FD value than healthy cases across all methods. Typical examples are presented in Figures 1 and 2. Figure 1 shows a clear separation between the normal with the lowest mean FD value (UTA0412) and a very acute index case (KSS0133).

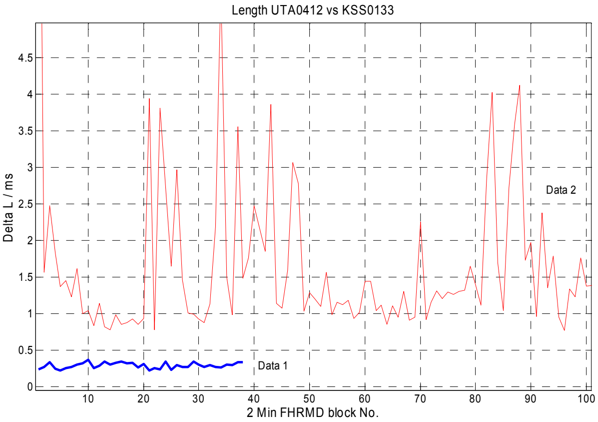

Figure 1. Showing the clear difference in FD between a normal and an index case. Data 1 - Index Case KSS0133; Data 2 - Normal Case UTA0412

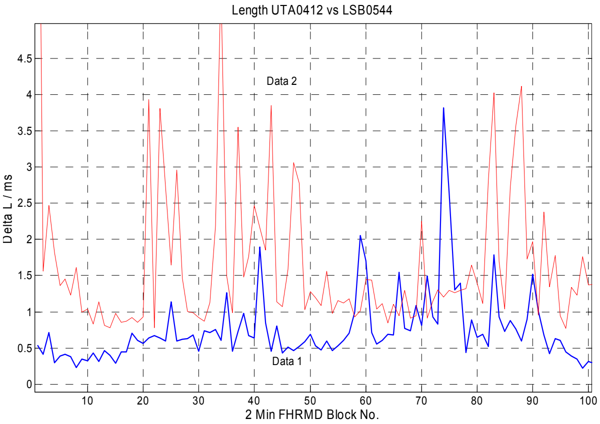

Figure 2. Showing the most challenging case where the index case most frequently overlaps into the normal region. Despite this, there is still a clear visible separation of in FD. Data 1 - Index Case LSB0544; Data 2 - Normal Case UTA0412

Figure 2 is the most challenging combination in the dataset, but even here it can be seen that although sporadic overlap occurs, it is not sustained and separation is still visibly clear. To manage sporadic changes in FD, a median filter with a window width of ten data points (20 minutes) was applied to the FD features. The risk of applying such a filter is an increase in latency before problems with a FHR trace are detected.

Figure 3 shows an example result of applying a median filter to the data used to produce Figure 2. The difference between healthy and compromised cases is shown more clearly allowing better distinction with minimal overlap.

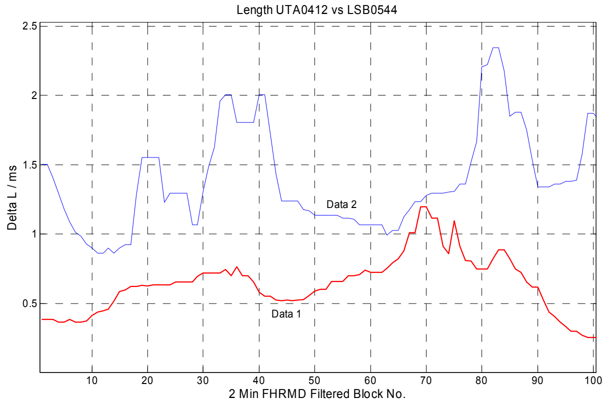

Figure 3. Effect of Median filter on figure 2. Data Data 1 - Index Case LSB0544; Data 2 – Normal Case UTA0412

To gain a clearer picture of how well each technique performs at separating cases, detection rates and contingency tables were produced for each technique. All median FD values for each case are tested against a range of threshold values. We decided to follow a (somewhat strict) policy for detection of abnormal cases: if any *one* median FD values falls below the threshold, then the case be flagged as a positive (abnormal), otherwise it is recorded as a false negative (normal). This was done as the clinical implication of missing an abnormal case (false negative) far outweighs false positive events. These results were used to calculate sensitivity and specificity values. The optimum threshold value is chosen such that the specificity value is obtained *given a 100% sensitivity*. These are shown for each technique in *Table 1*. Clearly, these preliminary results indicate that the modified Sevcik method has the best performance with an indicative specificity of just over 95% and the standard deviation as the worst at just less than 35%. Box counting could not be reliably applied to the data. When examined the filtered box counting dimension values for the index cases span range higher than that of the normal cases such that it will be impossible to achieve better than 0% specificity.

VI. DISCUSSION & CONCLUSION

Standard deviation was used as a benchmark figure. It can be seen in *Table 1* that this performs significantly worse than the length based FD estimators (Katz) with a specificity of just 34.78%. This is partly because such statistics ignore the ordering information in time series.

The basic Katz and Sevcik estimations give identical performance. This can be explained by examining the algorithms in section III.

Table 1. Contingency table for Analysis Techniques. Median Filter Window Length = 20 minutes

Analysis Technique	**Threshold Value**	**TP**	**FP**	**FN**	**TN**	**% Specificity**
SD	2.970	4	15	0	8	34.78
Katz	1.0000014	4	3	0	20	86.96
Length	1.200	4	3	0	20	86.96
Modified Sevcik	1.0360	4	1	0	22	95.65
Box Counting	1.396	4	23	0	0	0

Equation 6 shows that Katz's estimation is the ratio of the sum of the Euclidean distances through the set to the greatest Euclidean distance from one point to another. The FHR time series varies very little on the ordinate - typically the values range from 250-500ms, and far more significantly on the abscissa – 2 minute blocks give 120000ms. Thus the greatest Euclidean distance from one point to another will always be from the first to the last data value. This distance will always be extremely close to 120000 for a 2 minute block, and so it is advantageous in terms of processing complexity to assign it this value.

Sevcik includes the span on the abscissa by assigning $\varepsilon=1/(2N')$ where the 1 corresponds to the span. Both techniques use the ratio of the sum of the Euclidean distance through the set to its length. The difference in output value comes from the scaling technique used to relate this figure to the number of input data points.

When using fixed length blocks in all calculations it is not necessary to have this scaling technique as the output values can be directly compared. If not constrained to Fractal Dimension estimation then different indicators can be used giving the same performance. Perhaps the simplest of these is just subtracting the abscissa distance from the total Eulcidean distance. Data for this classification method is shown in *Table 1* as length technique.

The best performing algorithm featured is a Modified version of Sevcik's base algorithm. When applied without mapping Sevcik's algorithm provides identical performance to Katz's algorithm. This mapping has the effect of increasing the underlying separation between index and normal cases by spreading the cases over the ordinate. This favours high values (usually found in normal cases) and these receive the greatest increase in value from the mapping process further enhancing separation. It should be noted that the mapping process will lead to an artificially high estimation for fractal dimension; however this is of little concern as it is the application of the techniques to the classification of FHR which is the focus of this paper.

Box counting dimension gave the worst results in the sense that it could not be meaningfully applied to this data. It can be seen from its definition that the box counting method is a 'true' fractal dimension measure as it attempts to measure the variation of length with scale. As applied here it is not an effective marker for classifying FHR due to the non-fractal nature of FHR data. The box counting method measures the increase of line length with scale however there is no fine scale information in FHR data.

Approximations to fractal dimension provided by Katz's and Sevcik's algorithms are based on single scale measures give are not hindered by this drawback and therefore prove much better suited to this task. It is probable that the box counting dimension would prove more effective over some short specific edge lengths.

In conclusion it is possible to classify FHR data using fractal dimension estimates as applied here. Approximations to fractal dimension provided by Katz's and Sevcik's algorithms perform reasonably well. A clear distinction can only be guaranteed between compromised and normal cases in the longer term, thus rendering these markers less suitable for real time monitoring and clinical decision support. The modified Sevcik algorithm results in better performance at the expense of the accuracy of the FD estimate. With the exception of the box counting dimension, it is seen that these techniques offer significant improvement over simple Standard Deviation for the classification of FHR. No technique tested could completely separate all cases at all times within a case, and some overlap was observed. To overcome this, median filtering was used which inevitably introduces some latency. It remains to be shown if this is still clinically applicable.

It should be noted that the results presented here, though encouraging, have a relatively low statistical significance due to the limited numbers of cases used in the study. At the time of writing, the validation phase had not been completed. Our next step will be to validation these techniques against a much larger database of cases. The analysis of the results will be performed by clinical experts.

Further dimension estimation and other non–linear techniques will also be assessed to see whether different techniques or combinations of techniques offer any additional benefit.

REFERENCES

1. Computational Intelligence for Biopattern Analysis in Support of eHealthcare, *BIOPATTERN EU Network of Excellence. EU Contract 508803.*
2. Zhang, L., *Prenatal hypoxia and cardiac programming.* J Soc Gynecol Investig, 2005. **12**(1): p. 2-13.
3. Thornberg, E., et al., *Birth asphyxia: incidence, clinical course and outcome in a Swedish population.* Acta Paediatr, 1995. **84**(8): p. 927-32.
4. Amer-Wahlin, I., et al., *ST analysis of the fetal electrocardiogram during labor: Nordic observational multicenter study.* J Matern Fetal Neonatal Med, 2002. **12**(4): p. 260-6.
5. Harper, R.M., *The cerebral regulation of cardiovascular and respiratory functions.* Semin Pediatr Neurol, 1996. **3**(1): p. 13-22.
6. Echeverria, J.C., et al., *Does fractality in heart rate variability indicate the development of fetal neural processes?* Physics Letters A, 2004. **331**(3-4): p. 225-230.
7. Huikuri, H.V., T.H. Makikallio, and J. Perkiomaki, *Measurement of heart rate variability by methods based on nonlinear dynamics.* J Electrocardiol, 2003. **36 Suppl**: p. 95-9.
8. Yum, M.K. and J.H. Kim, *A very-short-term intermittency of fetal heart rates and developmental milestone.* Pediatr Res, 2003. **53**(6): p. 915-9.
9. Ott, E., *Chaos in Dynamical Systems*. 1993: Cambridge University Press.
10. Mandelbrot, B., *Fractals Form, Chance, and Dimension*. 1977, San Francisco: W. H. Freeman and Co.
11. Boshoff, H.F.V. *A fast box counting algorithm for determining the fractal dimension of sampled continuous functions*. IEEE Proceedings of the 1992 Symposium on Communications and Signal Processing,1992: p.43-48.
12. Katz's, M.J., *Fractals and the analysis of waveforms.* Comput Biol Med, 1988. **18**(3): p. 145-56.
13. Sevcik ,C. A Procedure to Estimate the Fractal Dimension of Waveforms. *Complexity International*, vol. 5, 1988.
14. Rosen, K.G. and N.J. Outram. *In Computational intelligence for biopattern analysis in support of eHealthcare*. in *1st EUROPEAN WORKSHOP ON THE Assessment of Diagnostic Performance*. 2004. Milano, Italy.
15. Siira, S.M. et al, Marked fetal acidosis and specific changes in power spectrum analysis of fetal heart rate variability recorded during the last hour of labour. BJOG, 2005; 112: 418-23.
16. Luttkus AK, et al. Fetal scalp pH and ST analysis of the fetal ECG as an adjunct to CTG. A multi-center, observational study. J Perinat Med 2004; 32(6):486-494