Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

A Comparative Study of Fetal Heart Rate Variability Analysis Techniques

Paul Hopkins, Nicholas Outram *Member, IEEE,* Nils Löfgren, Emmanuel C. Ifeachor, *Member, IEEE,* Karl G. Rosén

Abstract— **This study examines a novel methodology for continuous fetal heart rate variability (FHRV) assessment in a non-stationary intrapartum fetal heart rate (FHR). The specific aim was to investigate simple statistics, dimension estimates and entropy estimates as methods to discriminate situations of low FHRV related to non-reassuring fetal status or as a consequence of sedatives given to the mother. Using a t-test it is found that the dimension of the zero set and Sample entropy reveal a difference in mean distribution of significance >99%. Thus it may prove possible to build a discriminating system based on either one or a combination of these techniques.**

I. Introduction

The aim of intrapartum fetal surveillance is to identify fetuses at risk of hypoxia, thus enabling timely intervention to avoid an adverse outcome. The fetal electrocardiogram (FECG) provides FHR and FECG waveform patterns, the latter being objectively quantified by the STAN monitor (Neoventa Medical, Moelndal, Sweden). Visual assessment of fetal heart rate (FHR) patterns is still the norm and may be associated with substantial variation in the interpretation of FHR and decisions for intervention [1], [2].

Short term variations in the beat to beat intervals are thought to be the result of neurological control mechanisms and may contain information about neurological condition [3].

A healthy fetus is very capable in responding to the stress of labour [4]. However if adverse conditions persist or the fetus is vulnerable due to long-term deprivation of oxygen, there is the risk of neurological damage or even death.

Reassuring heart rate patterns such as accelerations and enhanced beat-to-beat variations indicate that the fetus is capable of reacting in a normal manner to intermittent stress caused by uterine contractions with variations observed during sleep states. These contractions cause large fluctuations in baseline FHR, limiting the use of techniques such as power spectrum analysis [5].

To address the issue of a non-stationary FHR, RR interval data are de-trended using a custom piecewise polynomial function approximation (the so-called primary FHR component). The deviation from the primary FHR component is termed the residual (Res) data. Decreasing or consistently reduced Res has been shown to accurately identify abnormal and low FHRV [6].

However, further refinements of the Res methodology may allow situations of low FHRV caused by drugs (sedatives; Pethidine) to be separated from those caused by adverse events (hypoxia, long term deprivation of oxygen). As a consequence, sedatives would disable the use of low FHRV as a specific marker. In order to include such cases there must be some fundamental difference in FHRV dynamics between those cases caused by adverse events and those due to medication and sleep states. If this exists then it may be possible to develop a more sensitive and specific technique/marker to successfully identify the at-risk cases whilst reducing false alarm rates. The focus of this study is the investigation into whether there are differences in the dynamics of a range of techniques with a view to developing a new marker.

II. Methodology

We have access to a high quality digitised database of STAN data (FECG and FHR) with known outcomes, including indications for operative delivery, cord acid base and neonatal data. For this study we use a selected dataset comprising four groups:

- Group I. - 5 “index” cases with adverse fetal outcome + no maternal medication and reduced FHRV.
- Group II. - 23 cases with normal outcome, normal FHRV, and no maternal sedatives administered.
- Group III. - 26 cases with normal outcome, visible FHR accelerations but where FHRV is suppressed due to administration of sedatives.
- Group IV. - 5 cases with normal outcome, *no* visible FHR accelerations and where FHRV was suppressed due to administration of sedatives.

In this study, we tested a range of different techniques to see if there are underlying differences in their ability to discriminate patterns of low FHRV recorded in the groups. The techniques used were:

Manuscript received April 24, 2006 resubmitted July 10, 2006. This work was supported in part by the Engineering and Physical Sciences Research Council.

P. Hopkins is with the School of Computing Communications and Electronics, University of Plymouth, Drake Circus, Plymouth, UK (phone: +44 (0)1752 232677; fax: (0)1752 232583; e-mail: paul.hopkins@plymouth.ac.uk).

N. Outram is with the School of Computing Communications and Electronics, University of Plymouth, Drake Circus, Plymouth, UK. e-mail: nicholas.outram@plymouth.ac.uk.

N. Löfgren is with Chalmers University of Technology, SE-412 96 Gothenburg, Sweden. e-mail: nilo@chalmers.se.

E. C. Ifeachor is with the School of Computing Communications and Electronics, University of Plymouth, Drake Circus, Plymouth, UK. e-mail: e.ifeachor@plymouth.ac.uk.

K. G. Rosén is with Neoventa Medical, Moelndal, Sweden. e-mail: kg.rosen@neoventa.com.

- Simple statistics: Standard Deviation and 95th percentile.
- Dimension estimates: Adapted Box Counting Dimension, Sevcik's estimate for Fractal Dimension, and the Dimension of Zero Set.
- Information theoretic techniques: Shannon entropy, Entropy of a First Order Markov Model, and Sample Entropy

For a mathematical description of the techniques used see appendix A.

Data acquisition and signal processing.

A STAN® S21 monitor recorded the fetal ECG during delivery with an intrauterine scalp electrode at 12 bits resolution, 500 Hz sampling rate. The STAN monitor software detected and stored R-peak interval times in digital recording files for further analysis.

To ensure data quality, pre-processing was applied as follows. Data was parsed for quality in two minute epochs. A rejection criterion of 4% missing samples from any two minute period is applied. Two minute blocks are selected such that total data loss is minimized.

To overcome the non-stationary nature of the FHR, a piecewise-polynomial curve fit approximation function was used to de-trend the FHR as follows: A window of FHR samples spanning a fixed 2 minute interval is selected; after linear interpolation of consecutive heart beats, the interpolated function is (evenly) re-sampled at 10Hz to which a piecewise polynomial is best fit (in a least-squares sense) – this produces the so-called primary FHR component; the residual data are generated by subtracting the primary FHR component from the re-sampled (10Hz) RR interval data.

The piecewise polynomial is derived from 5 consecutive polynomials of degree 6, each spanning 24s, with the added constraint that each polynomial shall have continuity with its neighbour(s) up to and including the first derivative (this can be extended to an arbitrary degree, but we found 1st derivative to be sufficient). This way, we obtain a smooth continuous function over the whole 2 minute interval with a maximum of 5 inflections per 24s and with some small dependence between the individual polynomial approximations.

The constraints are applied using the method of Lagrange multipliers. It can be shown that by formulating this with known polynomials (such as Legendre), we effectively derive new discrete and independent polynomials that more closely span the desired solution space [6], [7].

Data Selection

In this study we consider the most ambiguous scenarios presented by our data. For each technique applied, a contiguous 1 hour period was selected as follows:

- The single 1 hour period of FHR during a hypoxic episode (as identified by a clinical expert) from each case in Group I. that gives the lowest mean FHRV as defined by the specific technique.
- A single 1 hour period from each of the cases in Groups II, III and IV with the lowest mean FHRV as defined by the specific technique.

Thus for each technique, we test and compare the 1-hour blocks that are most ambiguous, i.e. have a mean closest to that of the Group I cases. We then test whether there are statistical differences between the cases in Group I and those in the three normal outcome groups (II, III and IV). Of particular interest are differences between Group I and those with reduced HRV (Groups III and IV).

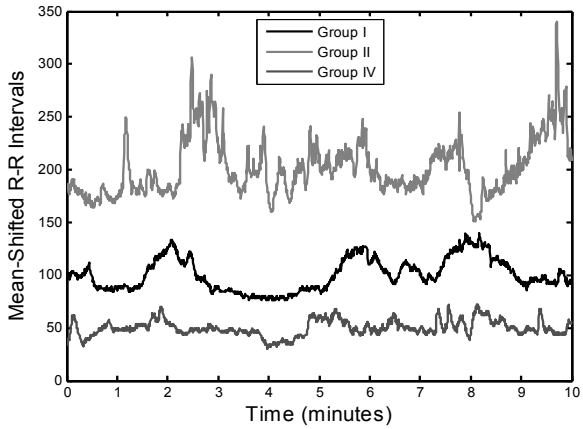

Fig. 1. Plot showing raw R-R interval data for Groups I, II and IV. Note data has been mean-shifted to clearly separate traces. The difference in variability between Group I and Group II can be clearly seen. Any difference in variability between Group I and Group IV is less pronounced.

Hypothesis Testing

For each technique, we test the Null hypothesis that there is no difference between the mean FHRV of Group I cases and groups II to IV. A Students t-test was applied with $\alpha<0.01$. Significance and confidence intervals are calculated at the given confidence level α.

III. Results

In order to reject the null hypothesis at the required confidence interval results of t-test must return p-values $< \alpha$ (0.01) see Table I overleaf. From the results in Table I, the two best performing techniques with statistically significant results are the Dimension of the zero set and Sample entropy.

Table I.
p-value results of t-test for each applied technique on selected 1 hour periods from each case - Group I cases against Groups II, III and IV. Dark shading represents rejection of the null hypothesis at confidence interval α<0.01

Technique	Group		
	II.	III.	IV.
Dimension Estimates			
Box Dim.	5.37E-29	2.05E-14	8.62E-02
Sevcik Dim.	1.77E-09	7.38E-04	0.104
Dim. Of Zero Set	2.35E-33	1.29E-24	8.68E-06
Simple Statistics			
Std Dev.	5.69E-36	1.05E-10	1.48E-02
95th Pct	7.24E-36	9.13E-10	2.62E-02
Entropy Measures			
Shannon	1.3E-08	0.658	0.819
F.O.M.M.	7.69E-26	7.45E-10	7.02E-05
Sample	2.58E-30	1.44E-10	4.39E-07

The Box Counting, Sevcik, standard deviation and 95th percentile techniques gave significant results for Groups II and III, but failed to give significant results for group IV. The Shannon entropy technique could only separate the Group I and Group II cases at the required confidence interval. This is discussed later in the paper.
Confidence interval plots (Fig. 2.) show underlying difference in mean value between groups.

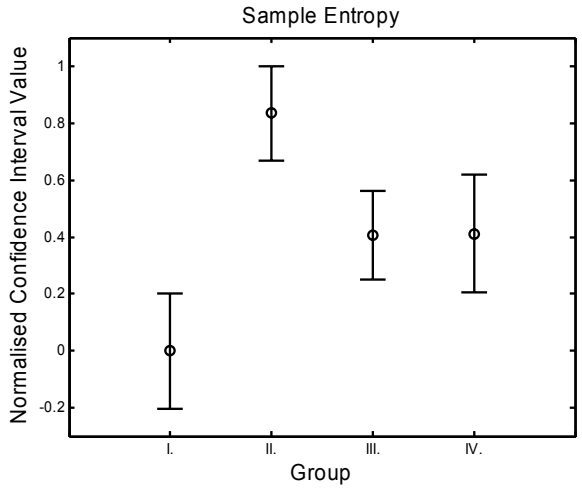

Fig. 2(a). Normalised plot of confidence interval (α=0.01) for difference in mean value between Group I and all groups using Sample Entropy as the applied technique.

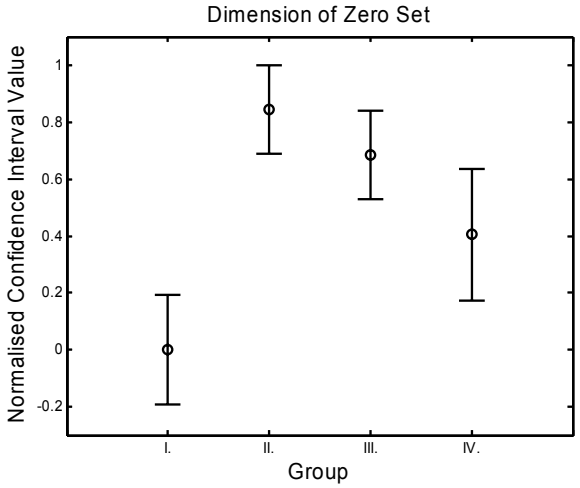

Fig. 2(b). Normalised plot of confidence interval (α=0.01) for difference in mean value between Group I. and all groups using the dimension of the Zero as the applied technique. Fig. 2 shows underlying difference in mean is much greater between Groups I and II than between Groups I and IV.

IV. Discussion

The techniques applied can be classified into three groups depending how they make use of data ordering or structure in producing a result.

- Simple statistics and Shannon entropy make no use of data ordering and produce results based on the PDF of the data.
- The Entropy of first-order Markov Model and Sevcik algorithms produce results based on sample-to-sample variation in the data. They use short term (immediate) sample ordering only.
- Techniques which make use of longer range sample order or structure are dimension of the Zero set, Box counting dimension, and Sample entropy.

We expected the best performance from techniques making greatest use of sample ordering information and indeed these techniques generally outperform the others. The exception is the box counting dimension which did not achieve significant results against Group IV cases. Performance of the box counting dimension may be altered by using different ranges of box edge lengths although we did not pursue this avenue.

Performance of simple statistical measures and Shannon Entropy is generally worse than that of the multi-scale dimension estimate techniques. This is probably because data structure (sample ordering) has no effect on these techniques. However performance was better than that some of the other dimension estimates which was unexpected. It is possible that discontinuities in the data have an adverse effect on the dimension estimates. The best performing techniques giving statistically significant differences between cases in Group I and the normal outcome groups

are Entropy of a first-order Markov model, Sample entropy and the Zero Set dimension. It is interesting to note that these all make use of structure information in the data. It also should be noted that these results do not necessarily mean that these techniques can be used to a strong discriminator/biomarker for adverse events in labour. They do indicate however there are likely to be some (non-visible) structural differences in FHR data between cases of reduced HRV due to hypoxia and artifacts due to the effects of administered medication. The findings will be used to further develop and investigate techniques to provide markers of adverse events during labour.

APPENDIX

Mathematical Description of Applied Techniques

A. Box counting dimension, D_b [8].

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

Where $N_{(\varepsilon)}$ is the number of boxes of edge length ε necessary to completely cover the set.

B. Sevcik Estimate for Fractal Dimension, D_s [9]

$$D_s = 1 + \frac{\ln(L)}{\ln(2N')} \tag{2}$$

Where L is the total length of the set, obtained by summing the distances between successive points and N' is the number of distances between points.

C. Dimension of the Zero Set, D_z [10]

$$D_z = -\left(\frac{\ln N(l)}{\ln l}\right) \tag{3}$$

Where $N_{(l)}$ is the number of lines of length l necessary to completely cover the set of zero crossing points after detrending.

D. Shannon entropy, H_{Sh} [11]

$$H_{Sh} = -\sum_{u=1}^{U} p(I_u)\log p(I_u) \tag{4}$$

$p(I_1)\ldots p(I_U)$, p_U is a discrete set of probabilities, which are estimated by counting the samples within a window falling in the disjoint amplitude intervals $I_1,..,I_U$, evenly distributed over -2 to 2 times the standard deviation. For $p = 0$, $p \log p$ is defined to be zero.

E. Sample entropy, H_S [12]

From a time series of N points N-m+1 series j of length m can be extracted if N>m. H_S is given by

$$H_S(r) = -\log\frac{B^{m+1}(r) - C^{m+1}_{N-m+1}(r)}{B^m(r)} \tag{5}$$

$$B^m(r) = \sum_i C_i^m(r) \tag{6}$$

$$C_i^m(r) = \frac{number\_of\_j : d(i,j) \le r}{N - m + 1} \tag{7}$$

$$d(i,j) = \max_{k=1,..,m}\left(\left|x(i+k-1) - x(j+k-1)\right|\right) \tag{8}$$

Where $d(i, j)$ is the maximum distance between an target sequence i of length m and all *other* j, $C_i^m(r)$ is the probability that $d(i, j)$ is smaller or equal to a given distance r where $i \neq j$, $B^m(r)$ is the mean of all $C_i^m(r)$ for a given value of r and $B^{m+1}(r)$ is the a result obtained by incrementing m by 1.

F. Entropy of first order Markov Model [13]

$$H_{MM} = -\sum_{u=1}^{U} p(I_u)\sum_{v=1}^{U} p(I_v \mid I_u)\log_2 p(I_v \mid I_u) \tag{9}$$

Where U is the alphabet of observed data values, $p(I_u)$ is the probability of occurrence of a specific value u given the previous value and $(I_v \mid I_u)$ represents the transition from value v to value u.

REFERENCES

[1] J.T. Parer and T. King, "*Fetal heart rate monitoring: is it salvageable?"* Am.J.Obstet.Gynecol., 2000,182:982-87.

[2] D.K. Donker, H.P van Geijn and A. Hasman, "*Interobserver variation in the assessment of fetal heart rate recordings,".* Eur J Obstet Gynecol Reprod Biol,1993,52(1):21-8.

[3] R.I. Kitney, and O. Rompelman, "*The Study of Heart-Rate Variability,"* Oxford University Press, 1980,

[4] K.R. Greene and K.G. Rosén, "*Intrapartum asphyxia,*" In Levene M et al (eds), Fetal and neonatal neurology and neurosurgery II, Churchill Livingstone, Edinburgh, 1995, pp 389-404.

[5] S.M. Siira *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 Apr,112(4):418-23.

[6] S. Blad, N. Outram, S. Arulkumaran, "*Continuous assessment of fetal heart rate variability during labor – Observations in healthy and asphyxiated fetuses*," "submitted for publication"

[7] http://www.earlylife.org.uk/lib/linalg/pwp.php

[8] H.F.V. Boshoff, "*A fast box counting algorithm for determining the fractal dimension of sampled continuous functions*. in *Communications and Signal Processing," 1992, COMSIG '92., Proceedings of the 1992 South African Symposium on*. 1992. Cape Town: IEEE.

[9] C.A. Sevcik, "*A Procedure to Estimate the Fractal Dimension of Waveforms*," Complexity International, vol. 5, 1988.

[10] R.F. Voss, "Fractals in nature: from characterisation to simulation, *The Science Of Fractal Images*," Springer-Verlag, pp 21-70. 1988.

[12] C. E. Shannon, *"A Mathematical Theory of Communication,"* The Bell System Technical Journal, vol. 27, pp. 379-423,623-656, 1948.

[13] J. S. Richman and J. R. Moorman, *"Physiological time-series analysis using approximate entropy and sample entropy,"* Am J Physiol Heart Circ Physiol, vol. 278, pp. H2039-2049, 2000.

[14] C. E. Shannon, *"Prediction and Entropy of Printed English,"* The Bell System Technical Journal, vol.30, pp 50-64, 1951.