The final publication is available at Springer via http://dx.doi.org/10.1007/s11517-014-1174-6

Recurring patterns in stationary intervals of abdominal uterine electromyograms during gestation

Luigi Yuri Di Marco[1,2], Costanzo Di Maria[3], Wing-Chiu Tong[1], Michael Taggart[1], Stephen Robson[1] and Philip Langley[1,4]

[1] Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne NE1 4LP, United Kingdom
[2] Department of Mechanical Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom
[3] Regional Medical Physics Dept. Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, United Kingdom
[4] School of Engineering, University of Hull, Hull HU6 7RX, United Kingdom

Address for Correspondence:

Luigi Yuri Di Marco, PhD

Department of Mechanical Engineering
The University of Sheffield
Mappin Street
Sheffield, S1 3JD
UK
Tel: +44 (0)114 2226074
l.dimarco@sheffield.ac.uk

Total number of words: 5533
Total number of word in Abstract: 200
Number of figures: 3
Number of tables: 5

Abstract

BACKGROUND: Abdominal uterine electromyograms (uEMG) studies have focused on uterine contractions to describe evolution of uterine activity and preterm birth (PTB) prediction. Stationary, non-contracting uEMG has not been studied.

AIM: To investigate recurring patterns in stationary uEMG, their relationship with gestation-age and PTB, and PTB predictivity.

METHODS: A public database of 300 (38 PTB) 3-channel (S1-S3) uEMG recordings of 30 minutes, collected between 22-35 weeks' gestation was used. Motion and labour-contraction free intervals in uEMG were identified as 5-minute weak-sense stationarity intervals in 268 (34 PTB) recordings. Sample entropy (SampEn), percentage recurrence (PR), percentage determinism (PD), entropy (ER) and maximum length (L_{MAX}) of recurrence were calculated and analysed according to time-to-delivery and PTB. Random time series were generated by random shuffle (RS) of actual data.

RESULTS: Recurrence was present in actual data ($p<0.001$) but not RS. In S3, PR ($p<0.005$), PD ($p<0.01$), ER ($p<0.005$), L_{MAX} ($p<0.05$) were higher, and SampEn lower ($p<0.005$) in PTB. Recurrence indices increased (all $p<0.001$) and SampEn decreased ($p<0.01$) with decreasing time-to-delivery suggesting increasingly regular and recurring patterns with gestation progression. All indices predicted PTB with AUC$\geq$0.62 ($p<0.05$).

CONCLUSION: Recurring patterns in stationary non-contracting uEMG were associated with time-to-delivery but were relatively poor predictors of PTB.

Keywords:

preterm birth; abdominal uterine electromyogram; recurrence plot; weak-sense stationarity; percentage determinism; entropy of recurrence; maximum length of recurrence; sample entropy

1. Introduction

The electrical properties of the abdominal uterine electromyogram (uEMG) after 24 weeks of pregnancy have been studied extensively over the past two decades, encouraged by the clinical interest for a non-invasive approach to uterine contraction monitoring and the prediction of time-to-delivery and preterm birth (PTB).

Abdominal uEMG signals have been studied in both time and frequency domains. Amplitude changes during contractions [8,19,22], time interval between contractions [19] and propagation properties [9, 21] have been shown to predict PTB. Nonlinear dynamics have also been successfully explored. Sample entropy, a measure of complexity (or conversely, regularity) of uEMG, was found to be a promising predictor of PTB in a study of 300 pregnancies [4]. Time reversibility, an indicator of nonlinearity in time series, was shown to discriminate between non-labour and labour contractions [9]. Devedeux and coworkers [2], in a review of uterine electromyography, showed that human myometrial (internal) and abdominal uEMG changes occur in phase with intrauterine pressure and exhibit similar spectra, including a slow wave in the frequency band 0.01–0.03 Hz likely caused by mechanical artefacts, and a fast wave. The latter was subdivided by Marque and colleagues [16] into a low frequency band (LFB: 0.2–0.45 Hz) associated with contractions during gestation, and a high frequency band (HFB: 0.8–3 Hz) associated with labour contractions. Accordingly, Garfield and colleagues [5] suggested that changes in the electrical properties of the uterus occur in the preparation for labour to make the myometrium more excitable and responsive in order to produce effective contractions capable of dilating the cervix. Maner and colleagues [15] have shown in a retrospective study on 99 subjects that peak frequency (f_P) of the power spectral density distribution (PSD) increased as the measurement-to-delivery interval decreased. Consistent results were reported by Garfield and colleagues [6] in a study of 50 subjects, in which those who delivered within 24 hours of recording had higher f_P than those who delivered later. The ability of spectral analysis to identify contractions leading to PTB was confirmed by Marque and colleagues [17]. However, uterine contractions occur infrequently preceding labor and may not be present during a typical abdominal uEMG

recording of 30 minutes duration. To our knowledge, wide-sense stationary (WSS) intervals, defined here as motion artefact-free and contraction-free intervals, have not been studied.

We hypothesise that contraction-free WSS intervals may exhibit deterministic, recurring patterns which could add further to our understanding of the mechanisms of myometrial preparation for labour and potentially contribute to the prediction of PTB. Recurrence analysis, originally presented by [1], has been applied to the dynamical assessment of physiological systems [1, 24]. The method is based on a graphical representation – termed recurrence plot (RP) – in which similar (recurring) sub-segments of a time series form diagonal lines on a state-space representation.

The aim of this study was threefold: *i)* to investigate the presence of recurring patterns in WSS intervals of uEMG; *ii)* to assess the relationship of recurrence indices with time-to-delivery and PTB; iii) to assess the potential contribution of these indices to PTB prediction.

2. Methods

2.1 Study population

A public database of 300 recordings of uterine electromyograms collected from 300 women between the 22^{nd} and 35^{th} week of gestation was used (http://www.physionet.org/pn6/tpehgdb) [7]. None of the women were in labour at the time of recording. For women delivering at term, the time-to-delivery was calculated as the difference between gestation duration and the time of recording. All recordings had a time-to-delivery between 2 and 22 weeks. This interval was divided into 4 equally sized non-overlapping bands of 5 weeks for subsequent analysis: Period 1, 2 to 7 weeks: Period 2, 7 to 12 weeks: Period 3 12 to 17 weeks: and Period 4, 17 to 22 weeks.

Each record consisted of 3 channels, recorded simultaneously from 4 electrodes placed at the corners of a square 7 cm apart and centred on the umbilicus as follows: 3.5 cm to the left and 3.5 cm above the umbilicus (E1); 3.5 cm to the right and 3.5 cm above the umbilicus (E2); 3.5 cm to the right and 3.5 cm below the umbilicus (E3); 3.5 cm to the left and 3.5 cm below the umbilicus (E4). The 3 channels were

acquired as: S1=E2–E1 (upper abdominal pair); S2=E2–E3; S3=E4–E3 (lower abdominal pair). Signals were digitized at 20 samples/s per channel with 16-bit resolution over a range of ±2.5 mV. Additional subject information was stored to a separate file for each record. Relevant subject information is summarized in Table 1. For each recording, three types of filtered data were available (0.3-3 Hz, 0.3-4 Hz, 0.08-4 Hz) as well as the unfiltered raw data. We chose to use the signals filtered in the band (0.08-4 Hz) in order to comply with previous studies [4, 22] while preserving the largest pass-band. To simplify the computational burden without loss of information, signals were down-sampled to F_S = 10 samples/s using a zero-phase anti-aliasing low-pass filter.

2.2 Identification of stationary intervals

Data stationarity was assessed in the "weak sense" (WSS), namely verifying time-invariance of the 1st and 2nd order statistical moments. The purpose was to identify contraction-free and motion artefact-free intervals of electrical activity.

The method used in this work was inspired by [14, 20]. Each signal of 30 minute duration (S1, S2, S3) was processed using a 5 minute sliding window (advancing by 1 s at each iteration) divided into 4 non-overlapping sub-windows ΔW_j (j=1,…,4) of 75 s each. Each window ΔW_j was treated as an independent observation of an N-dimensional ($N=75*F_S$) random variable ξ. Time invariance of the 1st order moment (mean) of ξ was assessed by the non-parametric Mann-Whitney test for univariate analysis of variance. Time invariance of the 2nd order moment (variance) was assessed by the non-parametric Levene test. For both tests a significance level of α=0.05 was used. Thus, p values greater than 0.05 indicated non-significant difference between samples (observations of ξ), which indicated stationarity of the 5-minute window being analysed. In the event of multiple instances of stationary excerpts in a recording, the first one in chronological order was retained for analysis. The choice for non-parametric tests was made to overcome the assumption of normally distributed data of parametric tests.

2.3 Assessment of nonlinearity

To test the hypothesis of the presence of nonlinear dynamics in WSS excerpts, the *time reversibility* test which a previous study [10] reported as the strongest predictor of nonlinearity in abdominal uEMG signals, was used. Briefly, this method is based on the null hypothesis that the time series of interest is originated by a Gaussian linear stochastic process (GSP). In this method, a GSP model is fitted to the original data, and N(=1000) surrogate time series are generated by the GSP process. A time reversibility metric (*Tr*) is calculated from the original data (Tr_{ORIG}) and from each surrogate time series (Tr_{SURR}). The difference $d_{Tr} = Tr_{ORIG} - Tr_{SURR}$ is calculated for each surrogate time series, and the distribution d_{Tr} is tested (Wilcoxon's rank test, 2-sided, $\alpha = 0.05$) against the null hypothesis that the original data –just like the surrogate– comes from a linear GSP process. In the presence of nonlinearity, the p value of the test statistics will exceed the significance level ($\alpha = 0.05$) and the null hypothesis will be rejected.

2.4 Random shuffle of uEMG

To quantify the presence of recurring patterns and predictability in the time series, the *random shuffle* (RS) method proposed by Hausdorff [11] was adopted. This method is based on the idea that shuffling the samples of a time series randomly does not alter the expected value, nor the variance of the time series, while it does remove any "memory effect", and hence any predictable or recurring pattern.

We quantified the presence of recurring and predictable patterns in the uEMG by comparing the recurrence and sample entropy indices for the original and the randomly shuffled data.

2.5 Recurrence analysis

To assess the presence of repeated (predictable) dynamics in uEMG the recurrence plot, an established phase-space based method, was used. Recurrence plot has successfully been applied to physiological time series to assess the regularity of the underlying system's dynamics [1, 18, 24]. An extensive presentation of the method can be found in [1,3,23]. Briefly, for a given time series the method is based on a graphical representation of the similarity of subintervals of fixed length *m*, known as the *embedding dimension*. On

a square matrix *R*, each point at coordinates *(i,j)* represents the similarity between two subintervals x_i and x_j both of length *m* (auto-recurrence analysis). Mathematically:

$$R_{i,j} = \begin{cases} 1, & \left\| x_i - x_j \right\| \le \varepsilon \\ 0, & otherwise \end{cases} \tag{1}$$

Where $\|\cdot\|$ represents the Euclidean distance operator, $x_i = (x_i \ \dots \ x_{i+m-1})^T$, and $(\cdot)^T$ is the transposed operator. This formulation assumes a delay constant $\tau = 1$ (1 sample), a typical value for discrete time series [23]. By construction, paths forming diagonals indicate repeated (recurrent) patterns.

This method offers the advantage of not requiring the definition of a model describing the system's dynamics, nor the assumption of data stationarity.

Previous studies [1, 24] have demonstrated that percentage recurrence (PR), percentage determinism (PD) and entropy of recurrence (ER) quantify many of the important characteristics of recurring dynamics in physiological signals. Mathematically these indices are expressed as follows:

$$\begin{aligned} PR &= \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j} \cdot 100 \\ PD &= \frac{\sum_{l=L_{MIN}}^{N} l \ H_l}{\sum_{i,j=1}^{N} R_{i,j}} \cdot 100 \\ ER &= -\sum_{l=L_{MIN}}^{N} p_l \log_2 p_l \\ p_l &= \frac{H_l}{\sum_{l=L_{MIN}}^{N} H_l} \end{aligned} \tag{2}$$

Where H_l is the proportion of diagonals of length l in the recurrence matrix R, p_l is the probability that a diagonal has length l. The index l ranges from L_{MIN} to the length N of the time series being analysed.

By definition, PR expresses the recurrence rate, indicating the percentage of recurring segments x_i in the time series; PD is a measure of the deterministic structure of the time series, quantifying the patterns forming diagonal lines (sustained recurrence) on the recurrence matrix R; and ER quantifies the distribution of diagonal lengths.

In the present study, in addition to the above indices, the maximum length of recurrence (L_{MAX}) was also considered:

$$L_{MAX} = \max_{k \in \{1,\ldots,N_D\}} \{l_k\} \tag{3}$$

Where N_D is the number of diagonal lines in the RP, and l_k is the length (number of points) of the kth diagonal line.

This was done to investigate the degree of nonlinearity in the system's dynamics. Indeed, periodic patterns in the time series (deterministic, linear dynamics) will result in long diagonal lines [23], whereas non-periodic patterns (non-deterministic, nonlinear dynamics) will result in shorter diagonals. In purely chaotic time series, diagonal lines do not appear (our results on random shuffle of the data time series are consistent with this property).

The embedding dimension m was empirically set to 2 s (20 samples), which was sufficiently high to satisfy the *false nearest neighbour* criterion [12] with tolerance threshold of 10%. The minimum length of diagonals (L_{MIN}) was set to 2 s (20 samples). The threshold ε was set case by case according to the criterion proposed by [1], namely 15% of the 95th percentile of the between-point distance distribution, which varied with each subject. An example illustrating the recurrence plot is shown in Figure 1.

2.6 Sample entropy

Sample entropy (SampEn) was also considered in this study. To calculate SampEn, the original time series is divided into sub-sequences of size m (embedding dimension). A metric is defined to quantify the

distance between any two sub-sequences of length *m*. The number of pattern matches (within a threshold distance *r*), is calculated for each value of *m*. Mathematically:

$$B^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\frac{B_i}{N-m-1}$$
$$SampEn\,(m,r,N) = -\ln\left[\frac{B^{m+1}(r)}{B^m(r)}\right] \quad (4)$$

Where B_i is the number of sub-segments of length *m* whose Euclidean distance does not exceed a fixed threshold *r*, and *N* is the length of the time series.

SampEn quantifies the probability of matching sub-sequences of length *m* to also match for length *m*+1. Hence, higher values of SampEn indicate higher complexity (or irregularity) of the time series. We used the values reported in a previous study [4] for *m* (=3) and *r* (=0.15 times the standard deviation of the time series). To allow comparison with [4] SampEn was calculated for the full 30 minute recording (SampEn30) as well as for the 5 minute stationary intervals.

2.7 Statistical analysis

Between-group differences were assessed by the Mann-Whitney U-test. *Post-hoc* analysis for multiple comparisons was done by the Tukey-Kramer test. All tests were two-tailed, with a significance level $\alpha = 0.05$. Receiver operating characteristic (ROC) analysis was computed to quantify the predictive ability of individual indices to predict PTB. Sensitivity (Se = TP/(TP+FN), TP = true positive (correctly classified preterm), FN = false negative), Specificity (Sp = TN/(TN+FP), TN = true negative, FP = false positive), accuracy (Ac = (TP+TN)/(TP+TN+FP+FN)) were calculated together with the positive (LR+ = Se/(1-Sp)) and the negative likelihood ratio (LR- = (1-Se)/Sp). For each index, the optimal threshold was identified as the point on the ROC curve with shortest Euclidean distance to the point corresponding to sensitivity and specificity of 1.

ROC analysis was done using SPSS Statistics™ v1.20 software (IBM Corp., Armonk, NY, USA). Other statistical analysis was done using MATLAB™ R2011b Statistical Toolbox™ software (The Mathworks, Natick, MA, USA).

3. Results

3.1 Group differences in clinical variables

Gestation at recording was similar in women who delivered at term (37-42 weeks of pregnancy) and preterm (<37 weeks of pregnancy) (Table 1). A history of "bleeding in 2nd trimester" was more common in the preterm birth group. However, the highly unbalanced distribution of observations with low numbers of positives ("Yes") for this parameter suggests caution in interpreting the χ^2 statistic.

3.2 Recurring patterns in baseline stationary intervals

A 5-minute stationary interval was found in the vast majority of signals. In particular, 268 recordings (32 of which were from the preterm group) had WSS in S1, 266 recordings (34 preterm) in S2, and 267 (31 preterm) in S3. Stationary intervals were detected 6 ± 6 minutes (mean±sd) into the recording.

According to the *time reversibility* test of nonlinearity, the majority (83%) of WSS excerpts were nonlinear in at least one channel (50% in S1, 68% in S2 and 56% in S3) confirming the suitability of nonlinear methods used in the analysis.

The absence of recurring patterns in time series was determined by a null value of the maximum length of recurrence (L_{MAX}). Conversely, recordings whose maximum length of recurrence was greater than zero had, by definition, recurring patterns. Recurring patterns were present in actual data but not in RS data, as shown in Table 2. RS data had higher SampEn (higher complexity) than original data (all $p<0.001$) showing that the random shuffling process effectively removed the predictability of the time series.

3.3 Nonlinear dynamics in women who delivered preterm and term

In the lower abdominal channel (S3) recurrence indices were higher in women who delivered preterm (all $p<0.05$), as shown in Table 3 and Figure 2. Consistently, SampEn for the stationary intervals was lower ($p<0.005$) in the preterm group. Indeed, SampEn is a measure of complexity of the time series, whereas recurrence indices quantify the presence of repeated behaviour; hence, an inverse relationship with SampEn is expected. SampEn calculated over the entire recording (SampEn30) in channel S3, had reduced significance compared to that from stationary intervals ($p<0.05$ vs. $p<0.005$).

3.4 Nonlinear dynamics in advancing pregnancy stage

The difference in time-to-delivery across periods was significant for all indices, for all channels (all $p<0.005$ in S1, S2; all $p<0.01$ in S3, Table 4) showing a trend for increasing values of recurrence indices (PR, PD, ER, L_{MAX}) and decreasing complexity index (SampEn), with shorter time to delivery. *Post-hoc* multiple group comparison showed a significant change ($p<0.05$, Table 4) comparing the earliest stages of gestation (Period 3, Period 4) to the later stage (Period 2), in all channels, for all indices except SampEn in S3 (Figure 3). Compared to SampEn, SampEn30 showed a reduced ability to discriminate between time-to-delivery strata (Table 4).

3.5 ROC Analysis for prediction of preterm birth

All indices from 5 minute stationary intervals in channel S3 were significant predictors of PTB (all $p<0.05$, Table 5). The area under the curve (AUC) was the highest for ER which also had the highest positive likelihood ratio (LR+), however the overlapping 95% CI of AUC indicate that none of the indices was a significantly better predictor than the others. Unlike SampEn, SampEn30 was unable to predict preterm birth ($p=0.08$).

4. Discussion

In this study the dynamics of baseline stationary intervals in abdominal uEMG recordings of gestation were analysed by nonlinear methods based on recurrence and regularity quantification. Unlike previous

works on abdominal uEMG focusing on pregnancy and labour contractions, the primary goal of the present work was to quantitatively assess the presence of recurring and predictable behaviour in stationary intervals outside of contractions, and their relationship to time-to-delivery interval and PTB prediction. In each channel, we found at least 266 (of 300) recordings that satisfied the *weak-sense stationarity* criterion, confirming the broad applicability of the method.

The presence of recurring patterns was shown (Table 2) in both preterm and at term delivery by comparing the data to a randomly shuffled version. Recurring patterns and regularity of stationary intervals were prominent in preterm compared to term delivery (Table 3), and increased with decreasing time-to-delivery (Table 4), suggesting increasingly regular and recurring patterns (increased degree of determinism) in later gestation and PTB. This hypothesis is consistent with the results presented by Hassan and colleagues [9] who showed more organised propagation of abdominal uEMG in labour contractions compared to non-labour contractions; and also with Garfield and colleagues' hypothesis [5] suggesting that changes in the electrical properties of the uterus occur in the preparation for labour. Our study suggests that these changes can be detected in relatively quiescent periods of uterine electrical activity.

Significant differences in term compared with preterm groups were only observed in the bipolar signal S3 (Table 3), which was measured on a horizontal electrode axis, in the lower region of the abdomen, closer to the cervical-isthmic section [4]. This suggests that the abdominal location and bipolar electrodes orientation may play an important role in determining the information conveyed by the electromyogram, which the nonlinear analysis aims to identify.

Sample entropy has been used in a previous study [4] on the same dataset, utilising the 30 minute recordings rather than 5 minute stationary excerpts used in our study. Interestingly, a similar discrimination ability between preterm and term birth was obtained in the above study, and also between earlier (<26 weeks) and later (≥26 weeks) gestation age, in spite of the longer observation window and the random presence and strength of contractions contained therein. As descriptive statistics (median and inter-quartile range) and AUC of sample entropy were not reported by [4], we calculated SampEn30 to

compare the predictive ability of the index in WSS intervals to that of the entire recording. Interestingly, only when calculated in stationary intervals was sample entropy a significant predictor of preterm birth (Table 5). Since stationary intervals were detected early in the recordings (on average within 6 minutes of the 30 minute recording) the analysis based on 5 minute stationary intervals has potential for reducing the required recording time compared to other analyses which is an important consideration in this population.

It should be noted that in our study all indices (including sample entropy) were calculated on down-sampled signals (10 samples/s vs. 20 samples/s). Although this may have influenced the calculated value of sample entropy (halving the sampling rate implies doubling the time span of the embedding dimension m), it is likely to be a minor factor as both SampEn and SampEn30 were calculated on the down-sampled signals, and the original signals were filtered in the frequency band 0.08 – 4 Hz.

Although individual indices from stationary intervals were all significant predictors of preterm birth (Table 5), their predictivity was relatively poor compared to previous studies using uEMG parameters derived from the analysis of contractions [13,17,19]. It should however be noted that in [13] and [19] the time of recording was much closer to delivery than in the dataset used in this study (7 and 14 days, respectively). Furthermore, a substantial difference in sample size between this study (N=300) and [13] (N=116), [17] (N=107) and [19] (N=87) may have played a role, as may have the difference in devices used for recording (sample rate, sample resolution), the number and position of electrodes, as well as environmental factors. As the dataset used in this study is public, investigation from other groups will prospectively allow comparison of results.

Further investigation of the relationship between gestational age (and time to delivery) and the nonlinear indices, considering term and preterm cases separately, is also warranted.

The public dataset used in this study only included 38 recordings from women who subsequently delivered preterm. This dataset of 300 recordings was a selected subset of more than 1200 recordings (not publicly available). Replication on a larger dataset from an unselected population is required to validate the proposed analysis of stationary intervals.

In conclusion, to the authors' knowledge this is the first study to quantitatively characterise stationary intervals in uEMG during gestation by recurrence analysis. The results provide quantitative evidence to support the hypothesis of changes occurring in the electrical properties of the uterus with advancing gestation, but recurrence indices in stationary intervals do not improve existing PTB prediction.

Acknowledgements

PL is funded by the National Institute for Health Research Newcastle Biomedical Research Centre based at Newcastle Hospitals Foundation Trust and Newcastle University. CM is supported by Research Capacity Funding from Newcastle upon Tyne NHS Foundation Trust. WT is funded by a Medical Research Council Bioinformatics Training Fellowship (grant no. G902091).

References

1. Censi F, Barbaro V, Bartolini P, Calcagnini G, Michelucci A, Gensini GF and Cerutti S (2000) Recurrent patterns of atrial depolarization during atrial fibrillation assessed by recurrence plot quantification. Ann Biomed Eng 28:61–70

2. Devedeux D, Marque C, Mansour S, Germain G and Duchêne J (1993) Uterine electromyography: a critical review. Am J Obstet Gynecol 169:1636–53

3. Eckmann JP, Kamphorst SO and Ruelle D (1987) Recurrence plots of dynamical systems. Europhys Lett 4:973–7

4. Fele-Žorž G, Kavšek G, Novak-Antolič Ž and Jager F (2008) A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups. Med Biol Eng Comp 46:911–922

5. Garfield RE, Saade G, Buhimschi C, Buhimschi I, Shi L, Shi SQ and Chwalisz K (1998) Control and assessment of the uterus and cervix during pregnancy and labour. Hum Reprod Update 4:673–95

6. Garfield RE, Maner WL, MacKay LB, Schlembach D and Saade GR (2005) Comparing uterine electromyography activity of antepartum patients versus term labor patients. Am J Obstet Gynecol 193:23–9

7. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101:e215–e220

8. Grgic O, Matijevic R, Kuna K (2012) Raised electrical uterine activity and shortened cervical length could predict preterm delivery in a low-risk population. Arch Gynecol Obstet 285:31–5

9. Hassan M, Alexandersson A, Terrien J, Muszynski C, Marque C and Karlsson B (2013) Better pregnancy monitoring using nonlinear propagation analysis of external uterine electromyography. IEEE Trans Biomed Eng 60:1160–6

10. Hassan M, Terrien J, Marque C, Karlsson B (2011) Comparison between approximate entropy, correntropy and time reversibility: application to uterine electromyogram signals. Med Eng Phys. 33:980–6.

11. Hausdorff JM (2009) Gait dynamics in Parkinson's disease: common and distinct behavior among stride length, gait variability, and fractal-like scaling. Chaos 19:026113

12. Kennel MB, Brown R and Abarbanel HD (1992) Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A 45:3403–3411

13. Lucovnik M, Maner WL, Chambliss LR, Blumrick R, Balducci J, Novak-Antolic Z, Garfield RE (2011) Noninvasive uterine electromyography for prediction of preterm delivery. Am J Obstet Gynecol 204:228.e1–10

14. Magagnin V, Bassani T, Bari V, Turiel M, Maestri R, Pinna GD, Porta A (2011) Non-stationarities significantly distort short-term spectral, symbolic and entropy heart rate variability indices. Physiol Meas 32:1775–1786

15. Maner WL, Garfield RE, Maul H, Olson G and Saade G (2003) Predicting term and pre-term delivery with transabdominal uterine electromyography. Obstet Gynecol 101:1254–60

16. Marque C, Duchene JM, Leclercq S, Panczer GS and Chaumont J (1986) Uterine EHG processing for obstetrical monitoring. IEEE Trans Biomed Eng 33:1182–7

17. Marque CK, Terrien J, Rihana S, Germain G (2007) Preterm labour detection by use of a biophysical marker: the uterine electrical activity. BMC Pregnancy Childbirth 7 Suppl 1:S5

18. Mohebbi M and Ghassemian H (2011) Prediction of paroxysmal atrial fibrillation using recurrence plot-based features of the RR-interval signal. Physiol Meas 32:1147–62

19. Most O, Langer O, Kerner R, David GB, Calderon I (2008) Can myometrial electrical activity identify patients in preterm labor? Am J Obstet Gynecol 199:378.e1–6

20. Porta A, D'Addio G, Guzzetti S, Lucini D, Pagani M (2004) Testing the Presence of Non Stationarities in Short Heart Rate Variability Series. Comp Cardiol 31:645–648

21. Rabotti C, Mischi M, Oei SG, Bergmans JW (2010) Noninvasive estimation of the electrohysterographic action-potential conduction velocity. IEEE Trans Biomed Eng 57:2178–87

22. Verdenik I, Pajntar M and Leskosek B (2001) Uterine electrical activity as predictor of pre-term birth in women with pre-term contractions. Eur J Obstet Gynecol Reprod Biol. 95:149–53

23. Webber CL Jr. (2012) Recurrence quantification of fractal structures. Front Physiol 3:382

24. Webber CL Jr, Zbilut JP (1994) Dynamical assessment of physiological systems and states using recurrence plot strategies. J Appl Physiol. 76:965–73

Figure Captions

Figure 1. Example of a recorded signal S3 from women who subsequently gave birth preterm (36.4 weeks) (left) and term (38.6 weeks) (right), with the corresponding recurrence plot. Preterm birth is associated with higher values for PR(6% vs. 2%), PD (40% vs. 23%), ER (4.83 vs. 4.26), and L_{MAX} (123 vs. 83), and lower SampEn (0.85 vs. 1.09).

Figure 2. Distribution of recurrence indices (PR, PD, ER, L_{MAX}) and sample entropy (SampEn) in preterm vs. term birth, for channels S1-S3. Shaded rectangles represent the inter-quartile range with median value, whiskers indicate 10^{th} and 90^{th} percentile, dots indicate 5^{th} and 95^{th} percentile. Significant between-group differences are indicated by horizontal brackets with *p value*.

Figure 3. Distribution of recurrence indices (PR, PD, ER, L_{MAX}) and sample entropy (SampEn) in term birth with respect to time to delivery, for channel S3. Shaded rectangles represent the inter-quartile range with median value, whiskers indicate 10^{th} and 90^{th} percentile, dots indicate 5^{th} and 95^{th} percentile. Significant between-group differences (*post-hoc* analysis) are indicated by horizontal brackets with *p value*.

Tables

Table 1. Subject Information.

		Term delivery (N = 262)		Preterm delivery (N = 38)		p value
		N(†)	Mean ± SD	N(†)	Mean ± SD	
Age [years]		217	29.4 ± 4.7	36	29.4 ± 4.8	N.S.
Time of recording [weeks]		262	26.8 ± 4.1	38	27.0 ± 3.8	N.S.
Gestation at delivery [weeks]		262	39.7 ± 1.1	38	34.4 ± 2.4	<0.001
Interval between recording and delivery [weeks]		262	12.9 ± 4.3	38	7.4 ± 4.1	<0.001
No. of previous deliveries		72	1.3 ± 0.6	12	1.2 ± 0.6	
No. of previous abortions		40	1.4 ± 0.7	7	1.4 ± 0.8	
Bleeding in 1st Trimester	Yes	28		4		N.S.
	No	213		27		
	Unknown	21		7		
Bleeding in 2nd Trimester	Yes	4		3		<0.01
	No	237		28		
	Unknown	21		7		
Smoker	Yes	12		1		N.S.
	No	157		18		
	Unknown	93		19		

(†) Number of recordings in which the given parameter was available

Table 2. Quantification of recurrence in preterm and term groups by comparison of original vs. random shuffled data. Values are Median(IQR)

	Signal	N_{TOT} (N_{PRE})	Index	Shuffled Data	Original Data	p value (Shuffled vs. Original)
Preterm	S1	268(32)	L_{MAX}	0(0)	62(34)	<0.001
			SampEn	2.42(0.12)	1.25(0.33)	<0.001
	S2	266(34)	L_{MAX}	0(0)	65(59)	<0.001
			SampEn	2.37(0.12)	0.92(0.61)	<0.001
	S3	267(31)	L_{MAX}	0(0)	61(53)	<0.001
			SampEn	2.38(0.13)	1.10(0.45)	<0.001
Term	S1	268(32)	L_{MAX}	0(0)	51(43)	<0.001
			SampEn	2.42(0.13)	1.25(0.36)	<0.001
	S2	266(34)	L_{MAX}	0(0)	63(48)	<0.001
			SampEn	2.41(0.12)	1.02(0.49)	<0.001
	S3	267(31)	L_{MAX}	0(0)	52(34)	<0.001
			SampEn	2.42(0.12)	1.26(0.37)	<0.001

N_{TOT}: total number of recordings with 5 min WSS excerpt; N_{PRE}: number of preterm delivery recordings with 5 min WSS excerpt

Table 3. Preterm vs. term effect on individual indices. Values are Median(IQR)

Signal	$N_{TOT}(N_{PRE})$	Index	Preterm	Term	p value
S1	268(32)	PR	0.9(1.9)	0.7(2.1)	N.S.
		PD	21.8(15.2)	18.1(17.2)	N.S.
		ER	4.22(0.80)	4.25(1.10)	N.S.
		L_{MAX}	62(34)	51(43)	N.S.
		SampEn	1.25(0.33)	1.25(0.36)	N.S.
		SampEn30	1.08(0.36)	1.11(0.44)	N.S.
S2	266(34)	PR	2.4(4.1)	1.8(3.1)	N.S.
		PD	23.6(21.2)	22.4(16.2)	N.S.
		ER	4.50(0.49)	4.38(0.49)	N.S.
		L_{MAX}	65(59)	63(48)	N.S.
		SampEn	0.92(0.61)	1.02(0.49)	N.S.
		SampEn30	0.79(0.39)	0.84(0.46)	N.S.
S3	267(31)	PR	2.0(4.6)	0.8(2.0)	<0.005
		PD	21.9(24.1)	18.3(13.3)	<0.01
		ER	4.45(0.67)	4.26(0.68)	<0.005
		L_{MAX}	61(53)	52(34)	<0.05
		SampEn	1.10(0.45)	1.26(0.37)	<0.005
		SampEn30	0.91(0.35)	1.05(0.43)	<0.05

N_{TOT}: total number of recordings with 5 min WSS excerpt; N_{PRE}: number of preterm delivery recordings with 5 min WSS excerpt.

Table 4. Time-to-delivery effect on individual indices (term birth recordings). Values are Median(IQR)

Signal	Index	Time-to-delivery [weeks]				p value*	Multiple Comparison			
		Period 1	Period 2	Period 3	Period 4		p value†	p value†	p value†	p value†
		2-7 weeks	7-12 weeks	12-17 weeks	17-22 weeks		Periods 1-3	Periods 1-4	Periods 2-3	Periods 2-4
S1 (N=12/93/84/47)	PR	1.6(4.2)	1.5(3.1)	0.4(1.0)	0.3(1.2)	<0.001	N.S.	<0.05	<0.05	<0.05
	PD	24.9(18.2)	22.4(19.7)	16.2(12.0)	14.0(13.0)	<0.001	N.S.	<0.05	<0.05	<0.05
	ER	4.51(0.41)	4.39(0.66)	4.12(1.05)	3.81(1.26)	<0.001	N.S.	<0.05	<0.05	<0.05
	L_{MAX}	81(64)	68(52)	47(31)	44(24)	<0.001	N.S.	<0.05	<0.05	<0.05
	SampEn	1.16(0.35)	1.16(0.37)	1.30(0.28)	1.33(0.43)	<0.005	N.S.	N.S.	<0.05	<0.05
	SampEn30	0.96(0.43)	1.06(0.43)	1.18(0.34)	1.15(0.46)	<0.01	N.S.	N.S.	<0.05	N.S.
S2 (N=10/88/90/44)	PR	3.6(16.4)	3.0(5.4)	1.2(1.9)	1.6(1.9)	<0.001	N.S.	N.S.	<0.05	<0.05
	PD	32.2(54.9)	29.9(26.3)	18.8(12.4)	20.6(10.0)	<0.001	<0.05	N.S.	<0.05	<0.05
	ER	4.51(1.66)	4.60(0.73)	4.29(0.40)	4.28(0.39)	<0.001	N.S.	N.S.	<0.05	<0.05
	L_{MAX}	86(162)	87(79)	55(38)	60(27)	<0.001	N.S.	N.S.	<0.05	<0.05
	SampEn	0.78(0.67)	0.85(0.47)	1.13(0.44)	1.06(0.41)	<0.001	N.S.	N.S.	<0.05	<0.05
	SampEn30	0.70(0.26)	0.73(0.52)	0.95(0.47)	0.89(0.44)	<0.001	N.S.	N.S.	<0.05	N.S.
S3 (N=13/94/83/46)	PR	2.3(5.5)	1.7(2.4)	0.5(1.2)	0.4(1.1)	<0.001	<0.05	<0.05	<0.05	<0.05
	PD	24.4(28.7)	21.6(13.8)	15.6(12.8)	15.6(9.4)	<0.001	<0.05	<0.05	<0.05	<0.05
	ER	4.50(0.81)	4.35(0.52)	4.21(0.95)	4.10(0.92)	<0.001	N.S.	<0.05	<0.05	<0.05
	L_{MAX}	87(75)	63(35)	47(28)	42(21)	<0.001	<0.05	<0.05	<0.05	<0.05
	SampEn	1.13(0.36)	1.17(0.37)	1.32(0.30)	1.36(0.36)	<0.01	N.S.	N.S.	<0.05	N.S.
	SampEn30	0.99(0.46)	0.99(0.42)	1.19(0.41)	1.06(0.41)	<0.01	N.S.	N.S.	<0.05	N.S.

N: Number of recording with WSS in Period 1 / Period 2 / Period 3 / Period 4.

* One-way analysis of variance (Mann-Whitney, $\alpha = 0.05$)

† *Post-hoc* analysis (multiple group comparison, Tukey-Kramer, $\alpha = 0.05$).

Table 5. Individual predictors of preterm birth (channel S3)

Index	AUC (95% CI)	p value	Optimal Threshold*	Se	Sp	Ac	LR+	LR-
PR	0.65(0.55, 0.75)	<0.01	1.93	0.52	0.70	0.68	1.74	0.69
PD	0.64 (0.54, 0.75)	<0.05	25	0.48	0.75	0.72	1.94	0.69
ER	0.67 (0.57, 0.77)	<0.005	4.44	0.55	0.74	0.69	2.07	0.61
L_{MAX}	0.62 (0.52, 0.73)	<0.05	72	0.45	0.75	0.72	1.81	0.73
SampEn	0.66 (0.56, 0.76)	<0.005	1.17	0.71	0.61	0.62	1.82	0.48
SampEn30	0.61 (0.50, 0.70)	N.S.	0.99	0.71	0.58	0.60	1.69	0.50

AUC: area under the curve; CI: confidence intervals; Se: sensitivity; Sp: specificity; Ac: accuracy; LR+: likelihood ratio positive; LR-: likelihood ratio negative;

*Calculated as the point on the ROC curve with shortest Euclidean distance to the point corresponding to sensitivity and specificity of 1