Liu Becky (Orcid ID: 0000-0002-4968-8294)
Thilaganathan Basky (Orcid ID: 0000-0002-5531-4301)
Bhide Amarnath (Orcid ID: 0000-0003-2393-7501)

Phase-rectified signal averaging: correlation between two monitors and relationship with short-term variation of fetal heart rate

B. Liu, B. Thilaganathan and A. Bhide

Fetal Medicine Unit, St George's University Hospitals NHS Foundation Trust, London, UK; Vascular Biology Research Centre, Molecular and Clinical Sciences Research Institute, St George's University of London, London, UK

Corresponding author: Dr A. Bhide

Fetal Medicine Unit, Department of Obstetrics and Gynaecology, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, UK

E-mail: abhide@sgul.ac.uk

Short title: Phase-rectified signal averaging in fetal ECG

Keywords: Phase-rectified signal averaging, short-term variation, fetal heart rate monitoring, non-invasive fetal electrocardiography, computerized cardiotocography, ambulatory monitoring

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.26192

CONTRIBUTION

What are the novel findings of this work?

Phase-rectified signal averaging (PRSA) obtained by using non-invasive fetal electrocardiography (NIFECG) and computerized cardiotocography (cCTG) are highly correlated. PRSA also has a strong linear relationship with short-term variation (STV), demonstrating its potential to assess fetal autonomic wellbeing.

What are the clinical implications of this work?

PRSA possess an inbuilt ability to eliminate noise, accounting for the dynamic NIFECG technology, thus generating higher accuracy and outputs than STV. PRSA may permit self-applied home fetal monitoring using NIFECG. Reliable remote fetal monitoring will enable increased fetal surveillance in high-risk women without increasing service demands.

ABSTRACT

Objective: To establish the correlation of phase-rectified signal averaging (PRSA) outputs between a novel self-applicable non-invasive fetal electrocardiography (NIFECG) monitor, and the computerized cardiotocograph (cCTG). A secondary objective is to evaluate the potential for assessing fetal wellbeing in the remote setting by assessing the relationship PRSA to short-term variation (STV).

Methods: This was a prospective observational study carried out in a London teaching hospital. Women with singleton pregnancies over 28+0 weeks' gestation attending hospital for cCTG assessment were recruited for concurrent cCTG and NIFECG monitoring for up to 60-minutes. Averaged accelerative and decelerative capacities (AAC and ADC) and STV were derived from both devices by post-processing. Signal filtration generated fully filtered (F-filtered) and partially filtered (P-filtered) results. Linear correlation, accuracy and precision analyses were performed to assess the relationship between monitors outputs, using varying anchor thresholds, and its association with STV.

Results: 306 concurrent cCTG and NIFECG traces were collected from 285 women. F-filtered NIFECG PRSA (eAAC/eADC) results were generated from 65% of cases where cCTG PRSA (cAAC/cADC) traces were generated. Correlations between cAAC/eAAC and cADC/eADC were strong (R=0.879, p<0.001, and R=0.895, p<0.001, respectively). NIFECG anchor detection reduced significantly with increasing signal loss, with large deviations from cCTG PRSA when <100 anchors were detected in a trace. Removing the anchor filters in NIFECG traces weakened the correlation (R=0.505, p<0.001, and R=560, p<0.001, respectively). Lowering the anchor threshold to 100 increased eAAC/eADC yield to 74%, whilst maintaining strong correlation with cCTG PRSA (R=0.839, p<0.001, and R=0.815, p<0.01, respectively). cAAC/cADC showed a very strong linear relationship to cCTG STV (R=0.928, p<0.001, and R=0.911, p<0.001, respectively). Similar findings were demonstrated with eAAC/eADC and cCTG STV (R=0.825, p<0.001, and R=0.827, p<0.001, respectively).

Conclusions: PRSA appears to be equivalent method of fetal assessment to STV, but due to its innate ability to eliminate artefact, PRSA is superior to STV in generating interpretable traces and trace accuracy with NIFECG. These findings raise the possibility of self-applied home or remote fetal heart-rate monitoring with automated reporting to enable increased surveillance in high-risk women without impacting service demand.

INTRODUCTION

Placental dysfunction remains a major health concern in pregnancy, manifesting in fetal growth restriction (FGR), chronic fetal hypoxemia and stillbirth[1,2]. Chronic hypoxemia can be assessed using methods such as fetal Doppler assessment, and fetal heart rate (FHR) analysis using cardiotocography (CTG). As fetal hypoxemia worsens, indices such as FHR variability declines[3,4]. The computerized CTG (cCTG) generates numerical outputs allowing standardized interpretation, minimizing inaccuracies from visual CTG assessments[5,6]. A key output is short-term variation (STV), which evaluates fetal autonomic activity through calculating the averaged mean 3.75 second (s) epochal differences[7]. STV has been shown to be a vital and reliable indicator of fetal hypoxia[6–9].

More recently, phase-rectified signal averaging (PRSA) has been proposed as a new method of fetal autonomic assessment. PRSA evaluates the speed of FHR change (oscillation) within quasi-periodicities, producing numerical averaged accelerative capacity (AAC) and averaged decelerative capacity (ADC). These indices allow the assessment of sympathetic and parasympathetic nervous systems on cardiac modulation as separate entities, whilst possessing the ability to account for non-stationary signals and eliminate noise[10–14]. Use of PRSA in adult electrocardiography (ECG) demonstrated the power of decelerative capacity in predicting mortality following myocardial infarction[15]. The significance of PRSA in autonomic assessment has prompted researchers to further evaluate its use in FHR measurements, with evidence suggesting that it may be superior to STV in detecting evolving fetal hypoxia[12–14,16].

Current fetal wellbeing assessments are hospital-based, limited by the availability of appointments and skilled personnel. The development of a reliable method of remote FHR monitoring with accurate automated outputs will enable increased fetal surveillance in high-risk pregnancies. Non-invasive fetal ECG (NIFECG) can be self-applied as it minimizes fetal-maternal heart rate confusion, is not limited by maternal adiposity and does not require placement in proximity to the fetal heart[17–20]. It is possible that previously reported challenges of susceptibility to artefacts in assessing STV may be overcome by PRSA's innate ability to eliminate noise[21,22]. The study objective is to compare PRSA values obtained by cCTG with those from a novel self-applied NIFECG device. A secondary objective is to evaluate the relationship between PRSA with STV values.

METHODS

This single-centered prospective cohort study was carried out at St George's University Hospitals NHS Foundation Trust, London. Recruitment took place from June 2021 to June 2022. Women with a singleton pregnancy over 28+0 gestation who presented requiring cCTG monitoring for any clinical indication were eligible, following written informed consent. Concurrent monitoring using both cCTG (Huntleigh Sonicaid FM800 Encore Fetal Monitor) and a novel self-applicable NIFECG (femom^{TM}), developed by Biorithm Pte Ltd., were performed for up to 60 minutes. This device aims to produce automated FHR outputs in the remote setting, with self-application and remote clinician assessment particularly in high-risk women. NIFECG data acquired were extracted in BioCapture recording files (bcrx), and exported following signal processing as comma separated value (csv) files, providing FHR values in each 0.25s epoch. Similar csv files were also exported from each concurrent cCTG trace. Full inclusion criteria, device information, and data acquisition methods are covered in detail in our study protocol[23].

Signal processing

NIFECG was sampled at a frequency of 500Hz. FHR generated through maternal R wave removal, fetal R wave signal enhancement, and RR interval calculation were in turn sampled at 4Hz, where each 0.25 second (s) epoch expressed an FHR value. cCTG sampling produced smoothed FHR values, which was also displayed in 0.25s epochs (4Hz). Signal acquisition in NIFECG is defined as FHR within a valid range. FHR outliers <30 or >240bpm are defined as signal loss. All FHR were in turn converted to fetal RR intervals (FRR), and expressed in milliseconds (ms).

STV computation

STV computation was as described by Dawes *et al*., where averaged differences of mean pulse intervals in 3.75s epochs were produced for each cCTG trace (cSTV)[5–7]. NIFECG STV (eSTV) follows a similar processing method. For optimal accuracy, three filters were applied at three points of the processing algorithm to remove traces or trace sections containing >50% signal loss, producing fully filtered (F-filtered) eSTV values. Partially filtered (P-filtered) eSTV values were generated without the application of these filters, instead only using an outlier filter. Our previous data on STV showed weak correlation between P-filtered eSTV and cSTV (R=0.337, p<0.001), but strong correlation when using F-filtered eSTV (R=0.911, p<0.001). For this reason, only F-filtered eSTV were used to compare PRSA values from the two monitors.

PRSA computation

PRSA processing followed the same steps and definitions as reported by Huhn *et al*, and were calculated for both NIFECG and cCTG (termed eAAC/eADC and cAAC/cADC respectively)[13]. T parameter is defined as time intervals, and using T=40 in our algorithm, 40 samples (4Hz = 4 samples per second, hence 10s windows) were averaged to create each T value. L is the length of surroundings, and is expressed in seconds. L=200 used in our algorithm means that a 200s window was extracted around each anchor (100s before and 100s after the anchor). S is the segment that defines the number of samples on either side of the final aligned anchor that was required to compute AAC/ADC. In our equation, S=40[13].

AAC values for each trace were generated using the following steps. *Step 1: Filter application* – each 40 sampling points were averaged to create T values (10s windows). Windows with >20% artefacts (or outliers) defined as FRR <250 or >2000ms (i.e., FHR >240 or <30bpm) were excluded. *Step 2: Anchor point definition* – T values decreasing from the previous T were assigned an anchor between the 2 values. *Step 3: Surrounding window definition* – 200s (L) around each anchor were selected. This can overlap with surrounding windows. *Step 4: Phase rectification* – all 200s windows were aligned with the anchors centered. *Step 5: Signal averaging* – an average waveform was created using all aligned waveforms. *Step 6: AAC quantification* – difference of the mean 39 samples before and mean 39 samples (S-1) after the aligned anchor. ADC is calculated in the same fashion, but anchors were defined between increasing T values in step 2[11–13].

AAC generated negative values due to the reduction in FRR, but absolute AAC values are reported in this paper for simplicity. Huhn *et al.* described removal of T values (averaged 10s windows) with >5% change from the preceding T for CTG PRSA computation[13], we adopted a slightly different approach of removing these 10s windows if >20% outliers (FRR <250 or >2000ms) were present. Two filters were applied in the processing of AAC and ADC to remove spurious and inaccurate results due to artefacts and signal loss. If the entire trace has >80% signal loss, and/or if there were <400 anchors detectable, eAAC or eADC values were not generated. In other words, only traces with low levels of signal loss were able to generate this fully filtered (F-filtered) PRSA output. Number of anchors detected for each NIFECG trace were recorded, and partially filtered (P-filtered) eAAC and eADC values were generated following removal of the anchor filter (all traces with any anchors detectable). Various anchor thresholds were used on the P-filtered dataset to define the best threshold to use for maximum data output whilst maintaining correlation with cCTG cAAC and cADC. As cCTG has lower levels of signal loss due to autocorrelation, high levels of anchors will be detected in each trace, therefore the analysis of anchor count was not performed for cCTG traces.

Statistical analysis

Descriptive data were presented as median and interquartile ranges (IQR) for continuous variables, and number and percentages for categorical variables. Linearity of AAC/ADC and STV values from the two devices produced by various methods of computation were established using Pearson's correlation coefficient, after confirming normality using the Kolmogorov-Smirnov test. Accuracy and precision analysis were carried out to assess the mean bias, precision (standard deviation), and 95% upper and lower limits of agreement (LoA), for each method of PRSA computation against cAAC/cADC. Statistical software package SPSS v28.0 (SPSS Inc., Chicago, IL, USA) was used for analysis. *P*-values < 0.05 were considered significant.

Ethical approval

Ethical approval was obtained from South-East Scotland Research Ethics Committee 02 (REC reference 19/SS/0109, IRAS ID 260032), and MHRA (CI/2020/0028).

RESULTS

Concurrent NIFECG and cCTG monitoring were undertaken in 285 women, with 306 traces collected. This study population was also used to investigate NIFECG signal loss and STV parameters, and the maternal and pregnancy characteristics are outlined in Table S1.

PRSA signal processing

Figure 1 displays the final AAC and ADC waveforms generated following signal averaging. As NIFECG is prone to artefacts and outliers, many did not fulfil the filtering requirements. Applying the PRSA anchor filter to NIFECG traces – 200/306 eAAC (65.4%) and 201/306 eADC (65.7%) traces produced PRSA values. Median eAAC and eADC were 5.5 (IQR 4.5 – 6.7) and 5.9 (IQR 4.6 – 7.4) ms respectively. Removal of these filters resulted in an increase in the outputs generated (255/306 eAAC (83.3%) and 256/306 eADC (83.7%)). In cCTG, a lower number of FHR outliers resulted in outputs from all traces. Median values of cAAC and cADC were 5.8 (IQR 4.6 – 7.0) and 6.1 (IQR 4.8 – 7.7) ms respectively. STV values were obtained from 46.4% of F-filtered and 98.4% of P-filtered NIFECG traces, and 100% of cCTG traces. Median and IQR for STV from the two monitors are displayed in Table 1.

PRSA agreement and anchor threshold definition

Table 2 demonstrates the agreement between F-filtered and P-filtered AAC/ADC between the two monitors. F-filtered eAAC and eADC were highly correlated with cAAC and cADC (R=0.879, p<0.001, and R=0.895, p<0.001, respectively). Mean bias and LoA for AAC are shown in the Bland-Altman plot in Figure 2. Conversely, P-filtered eAAC and eADC were less strongly correlated with cAAC and cADC (R=0.505, p<0.001, and R=560, p<0.001). Mean bias, precision and LoA also became more deviated from cAAC and cADC once the anchor filters were removed.

Using the Huhn method of removing 10s windows with >5% change from the previous T value, AAC and ADC from the two monitors were poorly correlated (cAAC vs eAAC: R=0.206, p<0.001, and cADC vs eADC: R=0.203, p<0.001). Relaxing our 20% outlier threshold of window removal to >30% outliers also resulted in a decline in AAC and ADC correlation (cAAC vs eAAC: R=0.206, p<0.001, and cADC vs eADC: R=0.413, p<0.001). Tightening the threshold to >10% outliers resulted in lower numbers of PRSA outputs generated (78.8% vs 83.7%).

Figure 3 illustrates all anchors detected in a high-quality signal trace section. The number of AAC anchors detected against percentage of signal loss in each trace is demonstrated in Figure 4. Leading up to 80% signal loss, very few anchors were detectable. AAC and ADC difference (eAAC – cAAC, and eADC – cADC) increased significantly when the number of

detected anchors reached below 100 (Figure 5 – AAC only). In order to increase the eAAC and eADC outputs whilst maintaining correlation with cAAC and cADC, different thresholds were applied at 100, 200, 300, and 400 (original filter) anchors detected. Agreement in AAC and ADC values in traces containing more than the aforementioned number of anchors were compared between the two monitors (Table 3). Linear correlation between eAAC/cAAC and eADC/cADC remained strong with reducing anchor thresholds. Comparing all traces with >100 anchors, higher number of traces yielded eAAC and eADC results (225/306 (73.5%) and 229/306 (74.8%) respectively), whilst correlation with cAAC and cADC showed little change (R=0.839, p<0.001, and R=0.815, p<0.01, respectively). Mean bias, precision and LoA also did not show significant change when anchor thresholds were reduced.

PRSA correlation with STV

There was a very strong correlation between cSTV vs cAAC and cADC (Figure 6 (AAC only): n=306, R=0.928, p<0.001, and R=0.911, p<0.001, respectively). F-filtered eAAC and eADC were also strongly correlated with cSTV (R=0.825, p<0.001, and R=0.827, p<0.001, respectively). Similarly, both eAAC/eADC and cAAC/cADC were strongly correlated with F-filtered eSTV (Table 4).

DISCUSSION

NIFECG can reliably produce PRSA values which are highly correlated with PRSA and STV values obtained using cCTG. Reducing anchor filters to 100 optimizes PRSA yield whilst maintaining concordance with cCTG and NIFECG PRSA values. Interpretable PRSA outputs were more frequently generated from F-filtered NIFECG. The strong correlation of PRSA with STV suggests that PRSA may accurately reflect fetal autonomic status.

PRSA signal processing

NIFECG samples at a frequency of 500Hz, is prone to artefact, electrical interference, and loss of R wave detection[21,22]. cCTG on the other hand, cannot sample at such high frequencies, and therefore uses autocorrelation to produce smoothed FHR at a frequency of 4Hz. FHR values from each 0.25s epoch will naturally be more fluctuant in NIFECG than cCTG. Higher T and S parameters will therefore allow more FHR samples to be averaged in each window, reducing the likelihood of high fluctuations in NIFECG. This is demonstrated in another study comparing PRSA outputs in CTG with NIFECG, using different parameters. 28 concurrent recordings from 9 FGR and 4 appropriate for gestation (AGA) fetuses showed that high correlation of AAC/ADC from the two monitors was seen only when T and S parameters were set at >40 (R>0.8), whilst minimal correlation was present when T and S were <5 (R<0.1)[24]. L parameter is simply the window size for anchor alignment, and does not influence the output.

Huhn's proposed method of removal of T values with >5% change from the preceding T was originally developed for cCTG. Removing 10s windows if >20% outliers are in line with Bauer's description of outlier removal in ECG[10]. The correlation demonstrates the superiority of this method, as it accounts for higher fluctuations in NIFECG, and that a 20% threshold allows maximum PRSA output whilst maintaining high correlation with cCTG. Other researchers have tested several algorithms and thresholds for T and S parameters, and applied these to cCTG and FECG outputs[11,14,24]. Several different proposals have been put forward for the best parameters, many for cCTG. Given the requirement for high T and S parameters in NIFECG, and the near-perfect correlation evident with cCTG STV vs cAAC/cADC using parameters set by Huhn et al.[13], T=40 and S=40 appear to be the most appropriate parameters to use in the assessment of fetal hypoxemia.

Clinical implications of PRSA

Previous comparisons have been made in FGR and AGA fetuses, all concluding that PRSA had superior diagnostic power than STV in FGR detection and/or predicting neonatal morbidity[11–13,25]. Lobmaier *et al*. demonstrated that in longitudinal cCTG recordings from 279

severe, early-onset FGR fetuses, cAAC/cADC declined 72 hours prior to elective birth, as opposed to STV, which significantly declined less than 48 hours before birth[16]. As our dataset captured a routine patient cohort, we did not include adequate numbers of FGR or adverse pregnancy outcomes to perform a similar comparison. The largest retrospective study was by Georgieva *et al.*, who applied the PRSA algorithm to 7568 intrapartum cCTG traces 30 minutes prior to delivery[14]. Increased cADC predicted acidemia at birth significantly better than STV (AUC 0.665 vs 0.606, p<0.001). Interestingly, the authors observed a weak correlation between STV and cAAC/cADC (R=0.29)[14]. This may be a result of the different parameters used for T and L (5 and 45) in their computation. It could also be due to the fact that STV is less likely to be accurate in active 2nd stage of labor, due to a high incidence of signal loss and decelerations. The mechanism of acidosis being more likely attributed to acute or sub-acute hypoxia as opposed to chronic hypoxia, thereby increasing FHR variability[26]. Indeed, other antepartum studies have however demonstrated stronger correlations between STV and cAAC/cADC[12,13]. Another intrapartum study compared PRSA and STV outputs obtained from cCTG in 227 cases of neonates born with acidemia, against 227 controls. This also found that ADC was significantly higher in those born with low umbilical artery pH, and was more predictive than STV (AUC 0.659 vs 0.566, p=0.013)[27].

Due to the strong linear relationship with STV as demonstrated in the current cohort, it would be reasonable to conclude that PRSA, with the appropriate computation, is equivalent to STV in assessing fetal autonomic status. Given the algorithm's innate ability to remove noise and artefacts, PRSA obtained by NIFECG can provide higher accuracy and increased outputs than STV. Although some authors have proposed a PRSA reference range, these were based on small sample sizes and has not been validated[12,25]. Until a validated reference standard exists, the linear relationship with STV may raise the possibility of setting PRSA thresholds according to its correlating STV value. Similarly, due to the difficulties encountered in accurate STV production from NIFECG found in our STV analysis, STV values can be generated through eAAC/eADC outputs.

Strengths and limitations

This study systematically evaluates PRSA outputs from two monitors and uses appropriate thresholds and algorithms which accounts specifically for the dynamic nature of NIFECG. We have further improved its output generation through pinpointing the exact cause for any inaccuracy. Its application to cCTG has resulted in high correlation with STV – the current FHR gold-standard in detection of hypoxia. Comparison with pregnancies complicated by placental insufficiency or those with adverse outcomes may have offered more insight into the diagnostic accuracy of PRSA, but due to our patient sample, this could not be performed.

Further research with a large, healthy cohort is required to establish and validate a reference standard.

Conclusion

PRSA appears to be an equivalent method of fetal wellbeing assessment than STV. With appropriate computation and signal processing of NIFECG to account for the difference in technology, it is highly correlated to cCTG PRSA and STV. Due to its in-built ability to eliminate artefacts, dynamic NIFECG can potentially replace the STV, which has demonstrated low accuracy and outputs in poor quality traces. These findings raise the potential for accurate, automated home FHR assessments using self-applied NIFECG, thereby improving fetal surveillance without increasing service demands in high-risk women.

ACKNOWLEDGEMENTS

We would like to acknowledge our commercial sponsors, Biorithm Pte Ltd, for funding this study, and the Bioengineering team for analysing the traces acquired and refining the processing algorithms. We would also like to thank the following midwives for identifying eligible women and helping with recruitment: Melissa Ogbomo, Maria Gutierrez-Vuong, Deborah Brown, Simone Schloss, Zainab Anidugbe, Yasmin Yussuff-Jakkari, Marta Hernandez, Ines Moya Sanchez, Inma Clavijo, and Manette Lindsay.

Disclosure

This study is funded by its commercial sponsors – Biorithm Pte Ltd. The sponsors have had no role in the study design, recruitment, data acquisition, or write up.

REFERENCES

1. Liu B, Nadeem U, Frick A, Alakaloko M, Bhide A, Thilaganathan B. Reducing health inequality in Black, Asian and other minority ethnic pregnant women: impact of first trimester combined screening for placental dysfunction on perinatal mortality. *BJOG* 2022; **129**: 1750–1756.

2. Poon LC, Shennan A, Hyett JA, Kapur A, Hadar E, Divakar H, McAuliffe F, da Silva Costa F, von Dadelszen P, McIntyre HD, Kihara AB, Di Renzo GC, Romero R, D'Alton M, Berghella V, Nicolaides KH, Hod M . The International Federation of Gynecology and Obstetrics (FIGO) initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int J Gynaecol Obstet* 2019; **145**: 1–33.

3. Ayres-de-Campos D, Spong CY, Chandraharan E. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *Int J Gynecol Obstet* 2015; **131**: 13–24.

4. Chandraharan E. Handbook of CTG Interpretation: From Patterns to Physiology. Cambridge, UK; New York: Cambridge University Press. https://books.google.co.uk/books?hl=en&lr=&id=qsAqDgAAQBAJ&oi=fnd&pg=PR11&dq=handbook+of+ctg+interpretation+chandraharan&ots=1zJnPzDdoS&sig=La68QTpKy2_j3TvP-nFONrdlWlk&redir_esc=y#v=onepage&q=handbook of ctg interpretation chandraharan&f=false. Published 2017. Accessed June 1, 2020.

5. Dawes GS, Lobb M, Moulden M, Redman CWG, Wheeler T. Antenatal cardiotocogram quality and interpretation using computers. *BJOG* 1992; **99**: 791–797.

6. Pardey J, Moulden M, Redman CWG. A computer system for the numerical analysis of nonstress tests. *Am J Obstet Gynecol* 2002; **186**: 1095–1103.

7. Street P, Dawes GS, Moulden M, Redman CWG. Short-term variation in abnormal antenatal fetal heart rate records. *Am J Obstet Gynecol* 1991; **165**: 515–523.

8. Bilardo CM, Wolf H, Stigter RH, Ville Y, Baez E, Visser GHA, Hecher K. Relationship between monitoring parameters and perinatal outcome in severe, early intrauterine growth restriction. *Ultrasound Obstet Gynecol* 2004; **23**: 119–125.

9. Ribbert LSM, Snijders RJM, Nicolaides KH, Visser GHA. Relation of fetal blood gases and data from computer-assisted analysis of fetal heart rate patterns in small for gestation fetuses. *Br J Obstet Gynaecol* 1991; **98**: 820–823.

10. Bauer A, Kantelhardt JW, Bunde A, Barthel P, Schneider R, Malik M, Schmidt G. Phase-rectified signal averaging detects quasi-periodicities in non-stationary data.

Phys A Stat Mech its Appl 2006; **364**: 423–434.

11. Stampalija T, Casati D, Montico M, Sassi R, Rivolta MW, Maggi V, Bauer A, Ferrazi E. Parameters influence on acceleration and deceleration capacity based on trans-abdominal ECG in early fetal growth restriction at different gestational age epochs. *Eur J Obstet Gynecol Reprod Biol* 2015; **188**: 104–112.

12. Lobmaier SM, Huhn EA, Pildner Von Steinburg S, Müller A, Schuster T, Ortiz JU, Schmidt G, Schneider KT. Phase-rectified signal averaging as a new method for surveillance of growth restricted fetuses. *J Matern Neonatal Med* 2012; **25**: 2523–2528.

13. Huhn EA, Lobmaier S, Fischer T, Schneider R, Bauer A, Schneider KT, Schmidt G. New computerized fetal heart rate analysis for surveillance of intrauterine growth restriction. *Prenat Diagn* 2011; **31**: 509–514.

14. Georgieva A, Papageorghiou A, Payne S, Moulden M, Redman C. Phase-rectified signal averaging for intrapartum electronic fetal heart rate monitoring is related to acidaemia at birth. *BJOG* 2014; **121**: 889–894.

15. Bauer A, Kantelhardt JW, Barthel P, Schneider R, Mäkikallio T, Ulm K, Hnatkova K, Schömig A, Huikuri H, Bunde A, Malik M, Schmidt G. Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study. *Lancet* 2006; **367**: 1674–1681.

16. Lobmaier SM, Mensing van Charante N, Ferrazzi E, Giussani DA, Shaw CJ, Müller A, Ortiz JU, Ostermayer E, Haller B, Prefumo F, Frusca T, Hecher K, Arabin B, Thilaganathan B, Papageorghiou AT, Bhide A, Martinelli P, Duvekot JJ, van Eyck J, Visser GHA, Schmidt G, Ganzevoort W, Lees CC, Schneider KTM; TRUFFLE investigators. Phase-rectified signal averaging method to predict perinatal outcome in infants with very preterm fetal growth restriction- a secondary analysis of TRUFFLE-trial. *Am J Obstet Gynecol* 2016; **215**: 630.e1–7.

17. Reinhard J, Hayes-Gill BR, Schiermeier S, Hatzmann H, Heinrich TM, Louwen F. Intrapartum heart rate ambiguity: A comparison of cardiotocogram and abdominal fetal electrocardiogram with maternal electrocardiogram. *Gynecol Obstet Invest* 2013; **75**: 101–108.

18. Graatsma EM, Miller J, Mulder EJH, Harman C, Baschat AA, Visser GHA. Maternal body mass index does not affect performance of fetal electrocardiography. *Am J Perinatol* 2010; **27**: 573–577.

19. Huhn EA, Müller MI, Meyer AH, Manegold-Brauer G, Holzgreve W, Hoesli I, Wilhelm FH. Quality predictors of abdominal fetal electrocardiography recording in antenatal ambulatory and bedside settings. *Fetal Diagn Ther* 2017; **41**: 283–292.

20. Graatsma EM, Jacod BC, Van Egmond LAJ, Mulder EJH, Visser GHA. Fetal electrocardiography: Feasibility of long-term fetal heart rate recordings. *BJOG* 2009; **116**: 334–337.

21. Sameni. A Review of fetal ECG signal processing issues and promising directions. *Open Pacing Electrophysiol Ther J* 2010; **3**: 4–20.

22. Clifford GD, Silva I, Behar J, Moody GB. Non-invasive fetal ECG analysis. *Physiol Meas* 2014; **35**: 1521–1536.

23. Liu B, Marler E, Thilaganathan B, Bhide A. Ambulatory antenatal fetal electrocardiography in high-risk pregnancies (AMBER): protocol for a pilot prospective cohort study. *BMJ Open* 2022 (In press).

24. Van Scheepen JAM, Koster MPH, Vasak B, Redman C, Franx A, Georgieva A. Effect of signal acquisition method on the fetal heart rate analysis with phase rectified signal averaging. *Physiol Meas* 2016; **37**: 2245–2259.

25. Graatsma EM, Mulder EJH, Vasak B, Lobmaier SM, Pildner von Steinburg S, Schneider KTM, Schmidt G, Visser GHA. Average acceleration and deceleration capacity of fetal heart rate in normal pregnancy and in pregnancies complicated by fetal growth restriction. *J Matern Neonatal Med* 2012; **25**: 2517–2522.

26. Bhide A, Johnson J, Rasanen J, Acharya G. Fetal heart rate variability with hypoxemia in an instrumented sheep model. *Ultrasound Obstet Gynecol* 2019; **54**: 786–790.

27. Weyrich J, Ortiz JU, Müller A, Schmidt G, Brambs CE, Graupner O, Kuschel B, Lobmaier SM. Intrapartum PRSA: a new method of predict fetal acidosis? A case-control study. *Arch Gynecol Obstet* 2020; **301**: 137–142.

FIGURE LEGENDS

Figure 1: Final averaged accelerative capacity (AAC) and decelerative capacity (ADC) waveforms generated after PRSA processing. FHR samples (4 samples per second) are plotted (x axis) against fetal RR intervals (FRR – y axis).

Figure 2: Bland-Altman plot demonstrating mean bias and limits of agreement (LoA) in AAC generated by cCTG and NIFECG. Bland-Altman plot with ADC outputs displayed a similar pattern.

Figure 3: Averaged accelerative capacity (AAC) and averaged decelerative capacity (ADC) anchors derived from an NIFECG trace section. Green and blue crosses annotate the AAC and ADC anchors respectively. FHR samples (4 samples per second) are plotted (x axis) against fetal RR intervals (FRR – y axis).

Figure 4: Scatterplot demonstrating the relationship between the number of anchors detected and signal loss for NIFECG AAC (eAAC). eADC displayed a similar pattern. Traces with >80% signal loss are automatically assigned 0 anchors.

Figure 5: Scatterplot demonstrating the relationship between AAC difference between the two monitors, and number of anchors detected. Logarithmic scale is used for AAC anchors on the X axis due to the high range of anchors. ADC showed a similar pattern.

Figure 6: Scatterplot showing correlation between cCTG AAC (cAAC) and cCTG STV (cSTV). cADC vs cSTV showed a similar pattern.

Table 1: Table demonstrating number of cCTG and NIFECG traces which generated PRSA and STV outputs, and their median values, using fully filtered (F-filtered) and partially filtered (P-filtered) processing algorithms.

	cCTG	**NIFECG F-filtered**	**NIFECG P-filtered**
PRSA			
Number of traces AAC	306 (100.0%)	200 (65.4%)	255 (83.3%)
Number of traces ADC	306 (100.0%)	201 (65.7%)	256 (83.7%)
AAC (ms)	5.8 (4.6 – 7.0)	5.5 (4.5 – 6.7)	5.6 (4.5 – 7.0)
ADC (ms)	6.1 (4.8 – 7.7)	5.9 (4.6 – 7.4)	5.9 (4.6 – 7.7)
STV			
Number of traces	306 (100.0%)	142 (46.4%)	301 (98.4%)
STV (ms)	9.9 (7.9 – 12.3)	9.2 (7.6 – 11.4)	10.7 (8.6 – 13.7)

Data are given as median (interquartile range) or number (%).

Table 2: Table showing linear correlation, accuracy, and precision analysis for F-filtered and P-filtered PRSA outputs, compared to cCTG PRSA.

	F-filtered cAAC/eAAC (n=200)	F-filtered cADC/eADC (n=201)	P-filtered cAAC/eAAC (n=255)	P-filtered cADC/eADC (n=256)
Pearson's R coefficient	0.879	0.895	0.505	0.560
Mean bias (ms)	-0.255	-0.283	0.191	0.062
Precision (ms)	0.848	0.924	2.488	2.732
Upper 95% LoA (ms)	1.407	1.528	5.067	5.417
Lower 95% LoA (ms)	-1.917	-2.039	-4.685	-5.293

cCTG outputs are termed cAAC or cADC, and NIFECG outputs eAAC and eADC. All p values <0.001.

Table 3: Table displaying P-filtered NIFECG PRSA using various anchor thresholds, and its correlation, mean bias, precision, and limits of agreement (LoA) against cCTG PRSA.

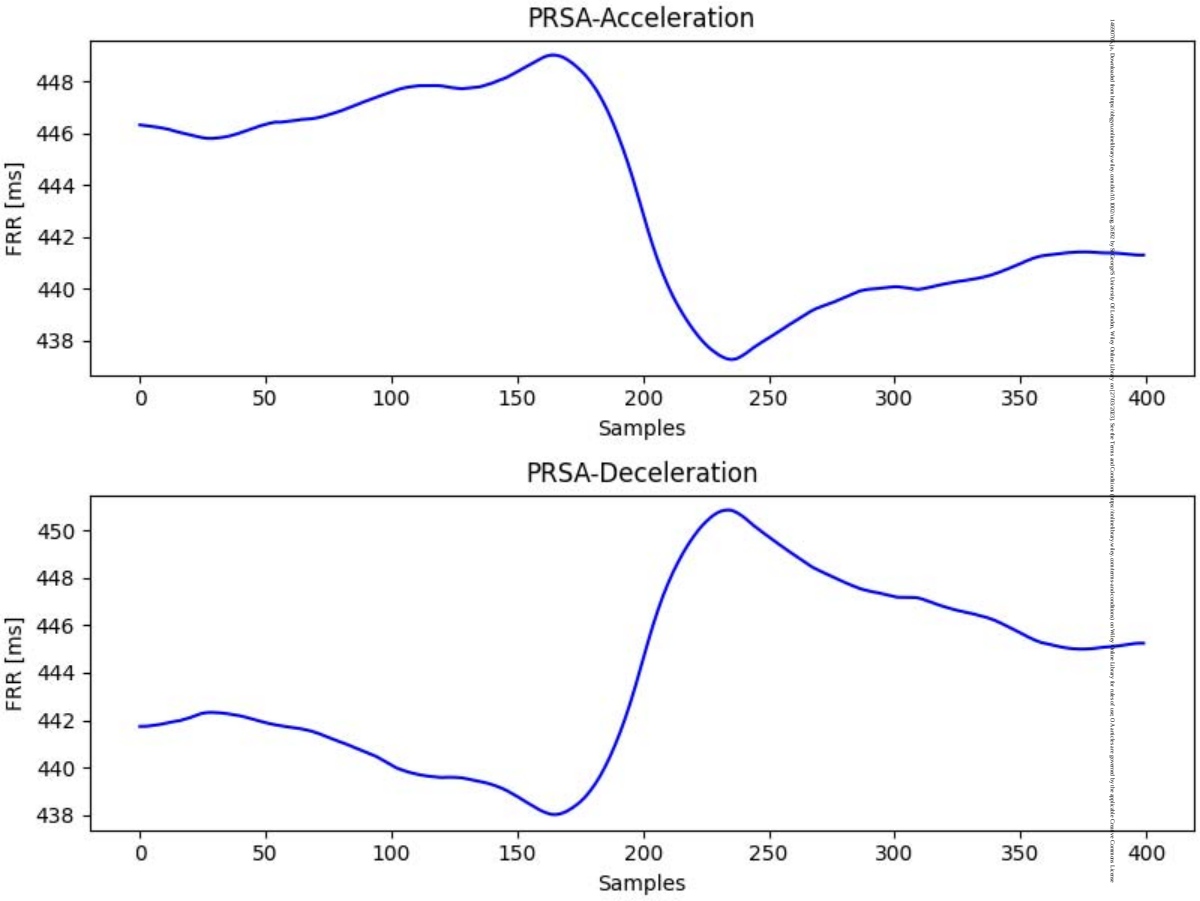
	eAAC/cAAC >400 anchors	eAAC/cAAC >300 anchors	eAAC/cAAC >200 anchors	eAAC/cAAC >100 anchors	eADC/cADC >400 anchors	eADC/cADC >300 anchors	eADC/cADC >200 anchors	eADC/cADC >100 anchors
Number of traces	200 (65.4%)	207 (67.6%)	215 (70.3%)	225 (73.5%)	201 (65.7%)	207 (67.6%)	216 (70.6%)	229 (74.8%)
Pearson's R	0.879	0.877	0.858	0.839	0.895	0.874	0.857	0.815
Mean bias (ms)	-0.255	-0.240	-0.231	-0.208	-0.283	-0.235	-0.200	-0.084
Precision (ms)	0.848	0.871	0.933	0.986	0.924	1.017	1.100	1.292
Upper 95% LoA (ms)	1.407	1.467	1.598	1.665	1.528	1.935	1.956	2.371
Lower 95% LoA (ms)	-1.917	-1.947	-2.060	-2.141	-2.039	-2.228	-2.356	-2.616

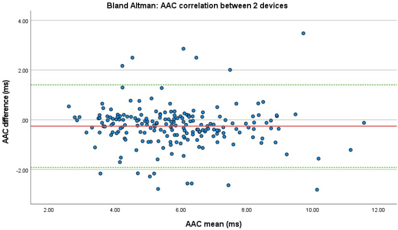
All p values <0.001.

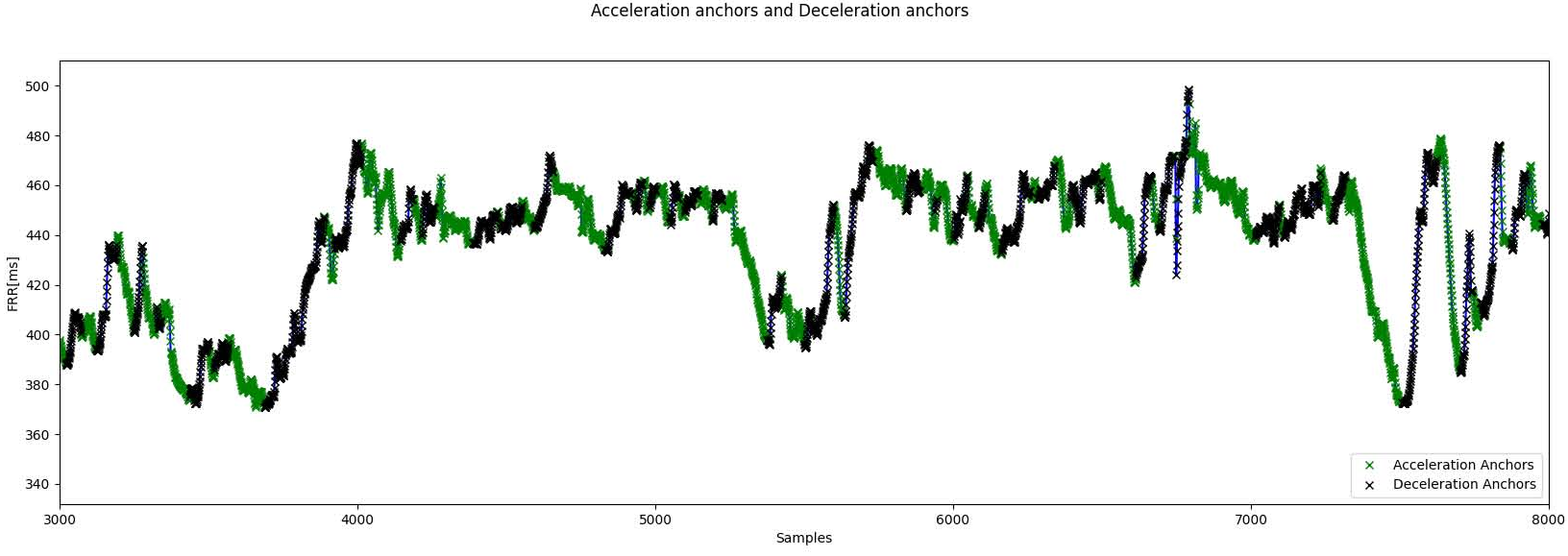
Table 4: Table displaying linear correlation between STV (cCTG and F-filtered NIFECG) and PRSA (cAAC/cADC and eAAC/eADC).

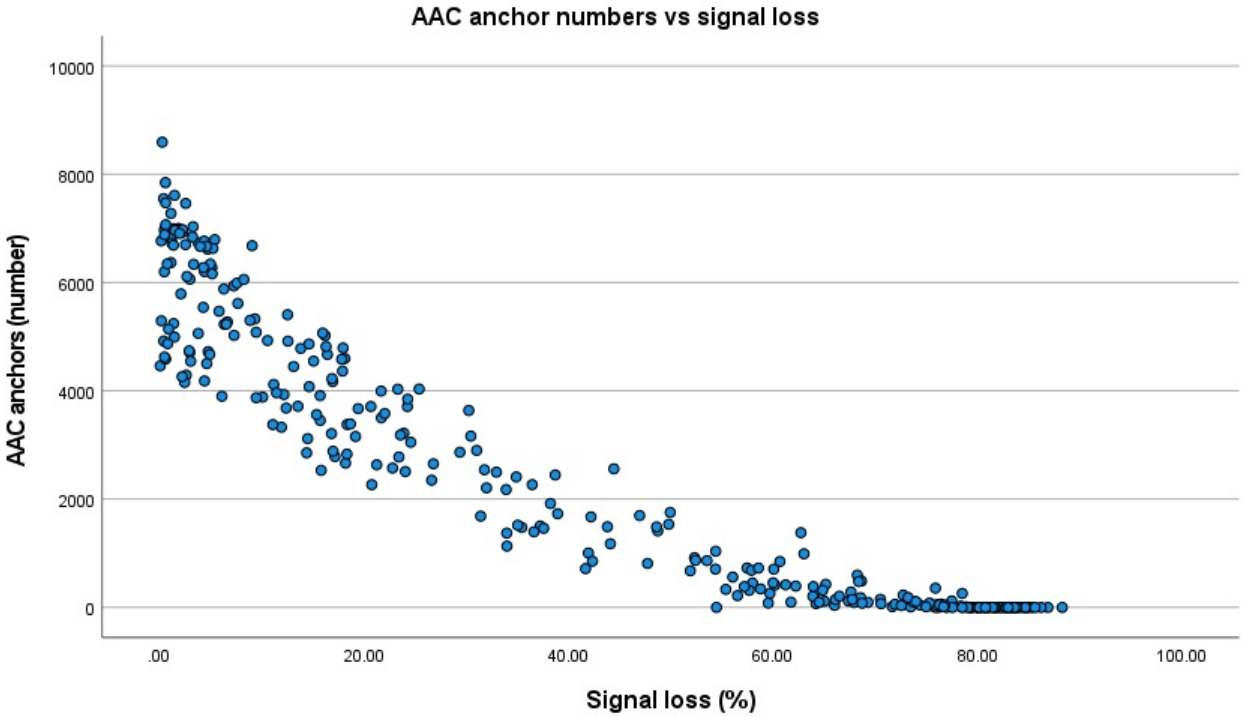
	cCTG STV (ms)	NIFECG STV (ms)
cAAC	0.928	0.838
cADC	0.911	0.834
eAAC	0.825	0.874
eADC	0.827	0.849

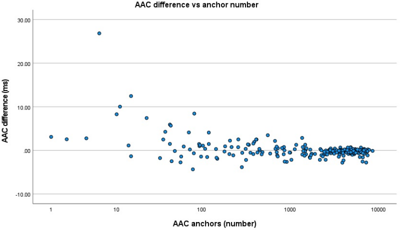
All p values <0.001.


PRSA-Acceleration
FRR [ms]
448
446
444
442
440
438
0
50
100
150
200
250
300
350
400
Samples
PRSA-Deceleration
FRR [ms]
450
448
446
444
442
440
438
0
50
100
150
200
250
300
350
400
Samples


Bland Altman: AAC correlation between 2 devices
4.00
2.00
.00
-2.00
AAC difference (ms)
2.00
4.00
6.00
8.00
10.00
12.00
AAC mean (ms)


Acceleration anchors and Deceleration anchors
FRR[ms]
Samples
Acceleration Anchors
Deceleration Anchors


AAC anchor numbers vs signal loss
AAC anchors (number)
10000
8000
6000
4000
2000
0
.00
20.00
40.00
60.00
80.00
100.00
Signal loss (%)


AAC difference vs anchor number
AAC difference (ms)
30.00
20.00
10.00
.00
-10.00
1
10
100
1000
10000
AAC anchors (number)