ORIGINAL RESEARCH ARTICLE

Relative accuracy of computerized intrapartum fetal heart rate pattern recognition by ultrasound and abdominal electrocardiogram detection

Barrie R. Hayes-Gill[1] | Terence R. P. Martin[2] | Chong Liu[1] | Wayne R. Cohen[3]

[1]Faculty of Engineering, Department of Electrical and Electronic Engineering, University of Nottingham, Nottingham, UK

[2]Monica Healthcare Ltd (retired), Nottingham, UK

[3]Department of Obstetrics and Gynecology, University of Arizona College of Medicine, Tucson, AZ, USA

Correspondence
Wayne R. Cohen, Department of Obstetrics and Gynecology, University of Arizona College of Medicine, 1501 N. Campbell, Tucson, AZ 85724, USA.
Email: cohenw@email.arizona.edu

Abstract

Introduction: Noninvasive fetal heart rate monitoring using transabdominal fetal electrocardiographic detection is now commercially available and has been demonstrated to be an effective alternative to traditional Doppler ultrasonographic techniques. Our objective in this study was to compare the results of computerized identification of fetal heart rate patterns generated by ultrasound-based and transabdominal fetal electrocardiogram-based techniques with simultaneously obtained fetal scalp electrode-derived heart rate information.

Material and methods: We applied an objective computer-based analysis for recognition of fetal heart rate patterns (Monica Decision Support) to data obtained simultaneously from a direct fetal scalp electrode, Doppler ultrasound, and the abdominal-fetal electrocardiogram techniques. This allowed us to compare over 145 hours of fetal heart rate patterns generated by the external devices with those derived from the scalp electrode in 30 term singleton uncomplicated pregnancies during labor. The direct fetal scalp electrode is considered to be the most accurate and reliable technique used in current clinical practice, and was, therefore, used as the standard for comparison. The program quantified the baseline heart rate, long- and short-term variability. It indicated when an acceleration or deceleration was present and whether it was large or small.

Results: Ultrasound was associated with significantly greater deviations from the fetal scalp electrode results than the abdominal fetal electrocardiogram technique in recognizing the correct baseline heart rate, its variability, and the presence of small and large accelerations and small decelerations. For large decelerations the two external methods were each not significantly different from the scalp electrode results.

Conclusions: Noninvasive fetal heart rate monitoring using maternal abdominal wall electrodes to detect fetal cardiac activity more reliably reproduced the computerized analysis of heart rate patterns derived from a direct fetal scalp electrode than did

traditional ultrasound-based monitoring. Abdominal-fetal electrocardiogram should, therefore, be considered a primary option for externally monitored patients.

KEYWORDS

Doppler ultrasound, electronic fetal monitoring, fetal heart rate monitoring, fetal heart rate variability, fetal scalp electrode, intrapartum fetal heart rate, transabdominal fetal electrocardiogram monitoring

1 | INTRODUCTION

Several investigators have shown that noninvasive transabdominal fetal electrocardiogram (ECG) monitoring is technically feasible and of clinical value.[1-8] We demonstrated previously that intrapartum fetal heart rate (FHR) detection using fetal ECG signals obtained by maternal abdominal surface electrodes (afECG) was more accurate and reliable than external ultrasound (US)-based monitoring when each external technique was compared with FHR data obtained simultaneously from a fetal scalp electrode (FSE).[8] That study assessed the ability of the two types of external monitoring to identify each fetal heartbeat detected by the FSE. It did not evaluate directly the accuracy of the external devices in identifying the FHR baseline, its variability, or the presence of accelerations and decelerations, central features of FHR pattern interpretation. Moreover, although a close correspondence was demonstrated in heart beat identification between afECG and FSE-derived data, the specific hypothesis that afECG is not inferior to traditional US-based monitoring in identifying FHR patterns (baseline FHR, variability, decelerations, and accelerations) was not tested.

To address this issue we applied a standardized computer-based approach for recognition of FHR patterns to data obtained simultaneously from a direct FSE, US, and the afECG technique. We used a computer-based approach to eliminate concerns about inter- and intra-observer variability. We compared FHR patterns generated by the external devices to those from the FSE. The FSE is considered the most accurate and reliable technique used in current practice, and was, therefore, used as the standard for comparison. The program quantified the baseline FHR, and long- and short-term variability, and it noted the presence and magnitude of accelerations and decelerations. Although most of these terms are used by clinicians to assess fetal well-being, for the purposes of this study we used them solely to describe the morphology of the FHR tracings, and not to infer anything about fetal condition.

2 | MATERIAL AND METHODS

We undertook a secondary analysis of data from a prospective 3-center trial designed to determine the performance of 2 modes of external FHR monitoring (afECG and US) compared with information obtained from an FSE.[8] For this report we examined data from 30 patients of one hospital (Queens Hospital Center, Jamaica, NY, USA) in the parent study. The hospital used the Model 50XM system

> **Key Message**
>
> Maternal transabdominal fetal electrocardiogram-based heart rate monitoring provided a more accurate depiction of heart rate variability, accelerations, and decelerations than ultrasound-based monitoring.

(Philips Healthcare, Andover, MA, USA) for standard FHR monitoring in all participants. The study design required the simultaneous use of 3 methods of FHR detection in each participant (afECG, US, and FSE). At the time of the study the afECG monitor (Model AN24; Monica Healthcare Ltd, Nottingham, UK) was being evaluated by the US Food and Drug Administration; it has since become commercially available.

In the parent investigation, 36 women entered the trial at the study hospital. We analyzed 30 of these in the first stage of labor; 24 were studied in the second stage. Six potential participants were excluded because the first-stage monitoring period for the 3 coincident tracings was <30 minutes (4 participants), or the US transducer malfunctioned, leaving an uninterpretable tracing (2 participants).

Each woman had a singleton term (≥37 weeks) pregnancy in cephalic presentation and was recruited early in spontaneous labor, or when she presented for labor induction. Each was monitored initially with the US technique. The tracing from the US monitor was available for clinical decision-making. The transducer position was adjusted as necessary by the attendant nurse every 20-30 minutes. The 5 afECG electrodes were applied to the maternal abdomen when it was determined that the US device was working appropriately. Data from the afECG were not visible to the care team.

An FSE was inserted later in labor in some patients if clinically indicated. The 30 women in whom simultaneous tracings from all 3 monitors were available for at least 30 minutes were the study participants.

FHR information from the US and the afECG monitors along with that from the FSE were transmitted to a bedside computer for analysis. No externally derived FHR data were visible to the care team once the FSE was applied, the US record was evaluated at least every 20–30 minutes by a research assistant who repositioned the transducer if necessary.

2.1 | Data collection and analysis

A software algorithm that can accurately identify FHR pattern characteristics was proposed by Dawes, Moulden, and Redman.[9-18] Their validated approach has been used in several studies of both antepartum and intrapartum FHR monitor tracings.[14,15,19-23] Engineers at Monica Healthcare, Ltd. created an algorithm (Monica Decision Support [MDS]) based on the principles established by Dawes and his colleagues (and as published by Dobbe et al)[24] and it was applied to assess the FHR patterns in our study. This algorithm was validated using a series of 87 hours of FHR data by the Divisie Perinatologie en Gynaecologie UMCU Utrecht (Utrecht, The Netherlands). This approach allowed us to examine the FHR patterns electronically in a uniform manner, avoiding the subjective observer bias that can confound consistent interpretation.

In order to generate the MDS parameters, FHR data were organized before analysis into 3.75-second epochs. The mean FHR ($FHR_{av(i)}$) in beats/minute (bpm) in each epoch was converted to a pulse interval in milliseconds for analysis, ie, $T_{av(i)} = 60\,000/FHR_{av(i)}$. The software then applied a rejection algorithm to each FHR pattern. $FHR_{av(i)}$ values <30 or >200 bpm and any $T_{av(i)}$ value >1.55 or <0.6 times the average of the preceding three $T_{av(i)}$ values were excluded. Also, any 30-minute window that contained <15 minutes of FHR data was not analyzed. After applying the rejection algorithm we identified the baseline FHR by deploying an exponential low-pass smoothing filter[24] having coefficients that resulted in a cut-off frequency of 1/600 Hz. The resulting time constant provided a slowly varying heart rate from which prominent changes (ie, accelerations and decelerations, the characteristics of which are described below) were removed, leaving behind the baseline FHR. Otherwise stated, the baseline FHR is the average FHR excluding accelerations and decelerations over a 30-minute period referred to below as a "frame." Our software also quantified long-term and short-term variability, and noted the presence of accelerations and decelerations.

Long-term variability (LTV) was assessed by determining over each minute the minute range, which is the absolute difference between the longest and shortest 3.75-second epoch pulse interval ($T_{av(i)}$), expressed in milliseconds over each sequential minute of baseline tracing. Minutes with more than 10 seconds of signal loss or that contained any part of a deceleration were excluded from LTV analysis. Then, the average LTV value over each 30-minute frame was calculated for analysis, thus generating the "mean minute range."

Short-term variability (STV) was measured by analyzing[25] the pulse interval of each 3.75-second epoch. The absolute difference from one epoch's pulse interval ($T_{av(i)}$) to the immediately following epoch's pulse interval ($T_{av(i+1)}$) expresses the short-term variation. Over 1 minute an average short-term variation is derived from these absolute differences. Then in each 30-minute frame the average STV was calculated and used in our analyses.

FHR accelerations and decelerations were identified by deviations from the baseline. A small acceleration was recorded when there was a rise in FHR from baseline of >10 bpm but ≤ 15 bpm that lasted >15 seconds; a large acceleration required a rise of more than 15 bpm and a duration of >15 seconds.

A small deceleration was noted when the FHR dropped between 10 and 20 beats on average for 60 seconds, or fell more than 20 beats from the baseline for 30 seconds. A large deceleration had a drop of at least 20 beats from baseline and a duration of at least 60 seconds or at least 40 beats for 30 seconds. The number of large and small decelerations and accelerations detected in each 30-minute block of the FHR tracing was compared between methods of FHR detection. A false-positive acceleration or deceleration was one that was noted on an external method but not by the FSE; conversely, a false-negative acceleration or deceleration was recognized by the FSE but not by the comparison external device.

2.2 | Statistical analyses

The FHR data from all three sources were synchronized to within 0.25 seconds before analysis. As noted above, the FHR values over each 3.75-second epoch were averaged into a single FHR value for that epoch; therefore, there were 16 FHR epoch values in every minute.

Using the MDS algorithm we extracted the baseline heart rate, STV, LTV, and the number of accelerations and decelerations in the simultaneous FHR recordings for each patient over all 30-minute frames in the first and second stages of labor. The median number of frame intervals analyzed per patient in stage I was 5, interquartile range 2–9, and in stage II was 3, interquartile range 2–4. In total for stage I and II there were 263 frames and 82 frames for afECG and 228 frames and 85 frames for US.

For each participant we identified all paired frame data points, ie, those in which the FSE standard and the comparison external monitor extracted an FHR analysis term. The paired data were used to determine differences between the distributions of data in the study groups. We calculated two sets of results for each participant: one compared the afECG against the FSE; the second compared the US method against the FSE. In each case the comparisons were made of the errors in the measurements, ie, the difference between the standard FSE value and the external value. Hence, the larger the error value, the greater the external device measurement deviated from that of the FSE. Because the distributions of these error measurements were not normal, we compared them using the nonparametric Wilcoxon signed-rank test.

Accuracy of the external monitoring modes in identifying each FHR parameter in comparison with the FSE standard was also assessed by Bland-Altman analysis, a standard approach to determining the degree of agreement between 2 measurements. We used the predicate standard FSE value on the horizontal axis and the difference between the comparison device and the FSE on the vertical axis.[25-27] In this manner a regression of the results from each pair of devices could be determined. This provided bias and limit of agreement values for each FHR data set comparison (FSE vs. afECG; FSE vs. US). The spread of values around the regression line is expressed mathematically by their root mean square error, limits of agreement, and bias.

A Z-test for difference of proportions was used to compare the rates of false-positive and false-negative accelerations and decelerations between monitoring modalities.

The statistical processing software deployed was IBM SPSS Statistics v. 24.0 (IBM Corp; Armonk, NY, USA). Matlab R2016A (The MathWorks, Natick, MA, USA) was used to create graphic Bland-Altman and box plots. All analyses were performed for all measured parameters on the complete pooled data from each monitoring modality.

2.3 | Ethical approval

The original study was approved by the Institutional Review Board of each participating institution, and conformed to the guidelines of the World Medical Association Declaration of Helsinki. For this report we examined data from 30 patients of one hospital (Queens Hospital Center, Jamaica, NY, USA) from the parent study. The Institutional Review Board of the Mount Sinai (NY) Medical Center approved the protocol (study #09-2213) on 5 January 2010.

3 | RESULTS

3.1 | Overall sample attributes

The gestational age (mean ± SD) of the participants was 39.4 ± 1.1 wk; maternal age was 25.9 ± 4.4 y; and body mass index was 32.1 ± 8.3 kg/m^2. All 30 study participants were monitored in the first stage of labor and 24 in the second stage. The monitoring duration per participant was 218 ± 187 min in the first stage and 91 ± 59 min in the second stage of labor.

TABLE 2 Distribution of errors

Parameter	afECG	US	P-value
Baseline FHR (bpm) stage I	0 (1) [263]	0 (1) [228]	0.007
Baseline FHR (bpm) stage II	0 (2) [82]	0 (6.3) [85]	0.036
STV stage I (ms)	0.07 (1.1) [263]	0.02 (2.8) [228]	0.012
STV stage II (ms)	0.08 (2.1) [82]	0.34 (5.9) [85]	0.002
LTV stage I (ms)	0.16 (3.9) [263]	1.63 (19.7) [228]	0.001
LTV stage II (ms)	−0.39 (15.6) [82]	5.7 (43.4) [85]	0.001
Small accelerations stage I (no./30 min)	0 (1) [263]	0 (1) [228]	0.008
Small accelerations stage II (no./30 min)	0 (3) [82]	0 (2) [85]	0.020
Large accelerations stage I (no./30 min)	0 (0) [263]	0 (1) [228]	0.001
Large accelerations stage II (no./30 min)	0 (1) [82]	0 (2) [85]	0.003
Small decelerations stage I (no./30 min)	0 (0) [263]	0 (1) [228]	0.001
Small decelerations stage II (no./30 min)	0 (1) [82]	0 (2) [85]	0.003
Large decelerations stage I (no./30 min)	0 (0) [263]	0 (0) [228]	0.015
Large decelerations stage II (no./30 min)	0 (0) [82]	0 (0.25) [85]	0.255

Note: All results are expressed as median (IQR) [n]. IQR, interquartile range; n, number of matched pairs in analysis; STV, short-term variability; LTV, long-term variability; P-values from Wilcoxon signed-rank test comparing the errors between the two monitoring methods. The error was the difference between the test device and the scalp electrode results.

TABLE 1 Summary of Bland-Altman analyses

Parameter	Device	n Stage I/ Stage II	LOA Stage I	LOA Stage II	Bias Stage I	Bias Stage II	RMSE Stage I	RMSE Stage II
Baseline rate (bpm)	afECG	263/82	−4.4, 4.3	−10.3, 12.1	−0.05	0.90	2.21	4.88
	US	228/85	−26.5, 19.1	−31.2, 22.8	−3.70	−4.20	11.40	12.50
STV (ms)	afECG	263/82	−3.9, 3.9	−9.7, 7.5	−0.02	−1.10	1.40	2.05
	US	228/85	−10.2, 15.0	−15.8, 22.1	2.40	3.10	6.20	8.80
LTV (ms)	afECG	263/82	−26.6, 23.3	−67.2, 46.4	−1.60	−10.40	8.70	11.90
	US	228/85	−70.5, 111.0	−93.4, 135.9	20.20	21.24	45.90	51.40
Small accelerations (No./30 min)	afECG	263/82	−2.8, 2.7	−5.5, 4.0	−0.04	−0.79	1.32	1.90
	US	228/85	−3.8, 4.9	−6.6, 8.5	0.56	0.96	2.15	3.80
Large accelerations (No./30 min)	afECG	263/82	−2.7, 2.5	−5.3, 3.5	−0.06	−0.89	1.17	1.35
	US	228/85	−3.6, 4.8	−6.2, 8.3	0.60	1.05	2.13	3.70
Small decelerations (No./30 min)	afECG	263/82	−1.9, 2.0	−3.7, 3.1	0.03	−0.33	0.94	1.40
	US	228/85	−3.0, 3.9	−4.1, 5.0	0.48	0.46	1.66	1.99
Large decelerations (No./30 min)	afECG	263/82	−0.6, 0.6	−1.8, 1.2	0.00	−0.27	0.31	0.60
	US	228/85	−1.5, 1.7	−2.6, 2.2	0.13	−0.19	0.79	0.90

Abbreviations: RMSE, root mean square error; afECG, abdominal-fetal ECG device; US, Doppler ultrasound device; STV, short-term variability; LTV, long-term variability; n, number of matched pairs stage I/stage II; LOA, limits of Agreement.
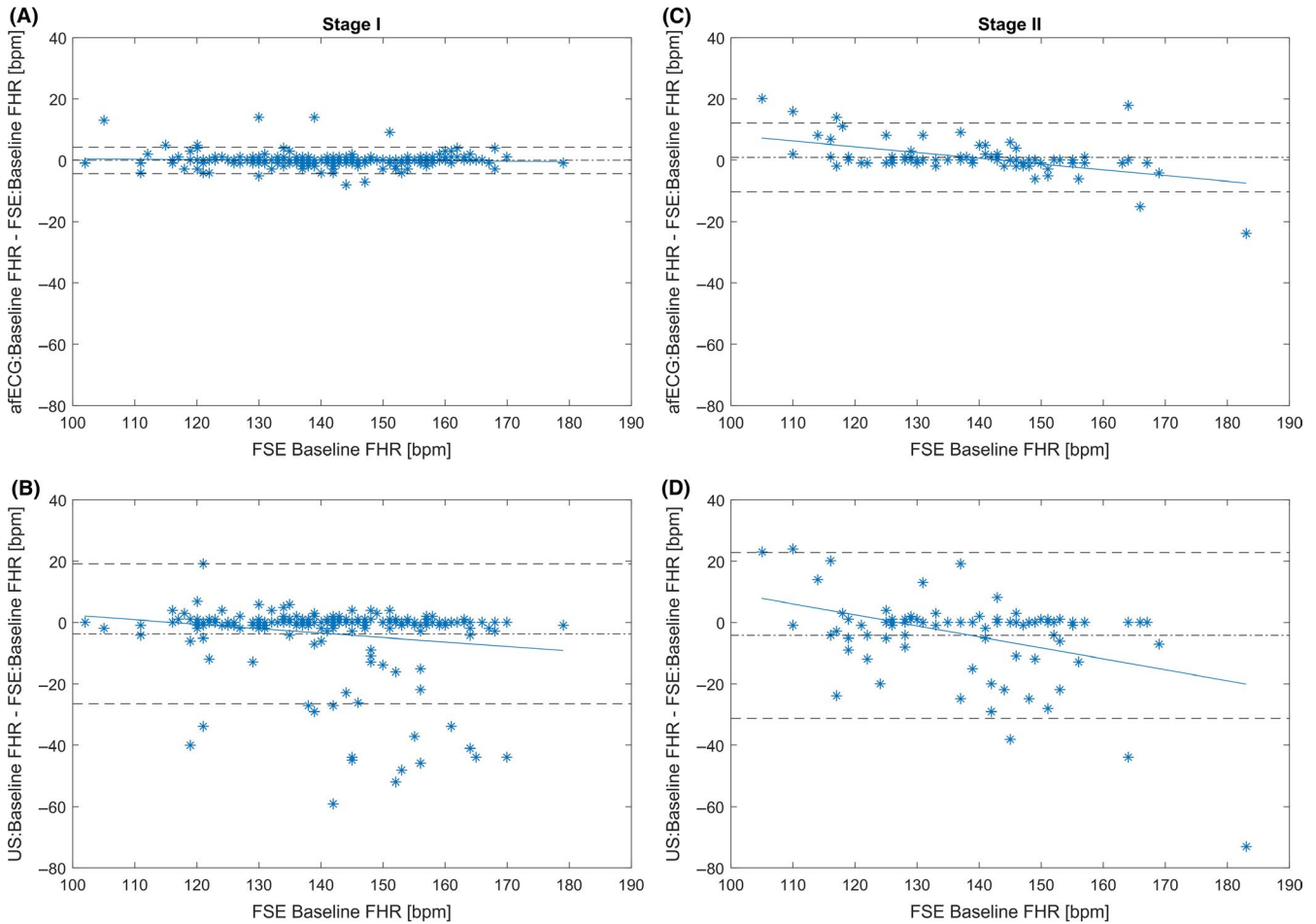
FIGURE 1 (A-D) Bland-Altman plots for baseline fetal heart rate (FHR) in stage I of labor comparing agreement of abdominal-fetal electrocardiogram (afECG) method with fetal scalp electrode (FSE) method (A) and ultrasound (US) with FSE (B). Stage II results are also shown for afECG (C) and US (D). In both stages the afECG displayed better agreement with the FSE than did US [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 False-positive and false-negative periodic event rates

Modality	Parameter	False-negative rate (number/30 min)		False-positive rate (number/30 min)	
		Stage I	Stage II	Stage I	Stage II
afECG n = 263/82	Small accelerations	0.44**	1.18*/**	0.40**	0.39
	Large accelerations	0.35**	1.12*/**	0.29**	0.23
	Small decelerations	0.25	0.70*/**	0.28**	0.37
	Large decelerations	0.04	0.32*	0.04**	0.05
US n = 228/85	Small accelerations	0.35	0.67*	0.91	1.64*
	Large accelerations	0.20	0.48*	0.80	1.53*
	Small decelerations	0.25	0.55*	0.74	1.01*
	Large decelerations	0.07	0.40*	0.20	0.21

Note: n, number of 30-min frames analyzed in stage I/stage II; afECG, abdominal-fetal ECG method; US, Doppler ultrasound method. Data are the mean number of false-positive or false-negative accelerations or decelerations per 30-minute frame. A false-positive event is one that occurred in the test device but not in the scalp electrode data; a false-negative event was seen in the scalp electrode but not in the test device.

*Stage II significantly different from stage I using Z-test.

**Significantly different from US for same parameter using Z-test.
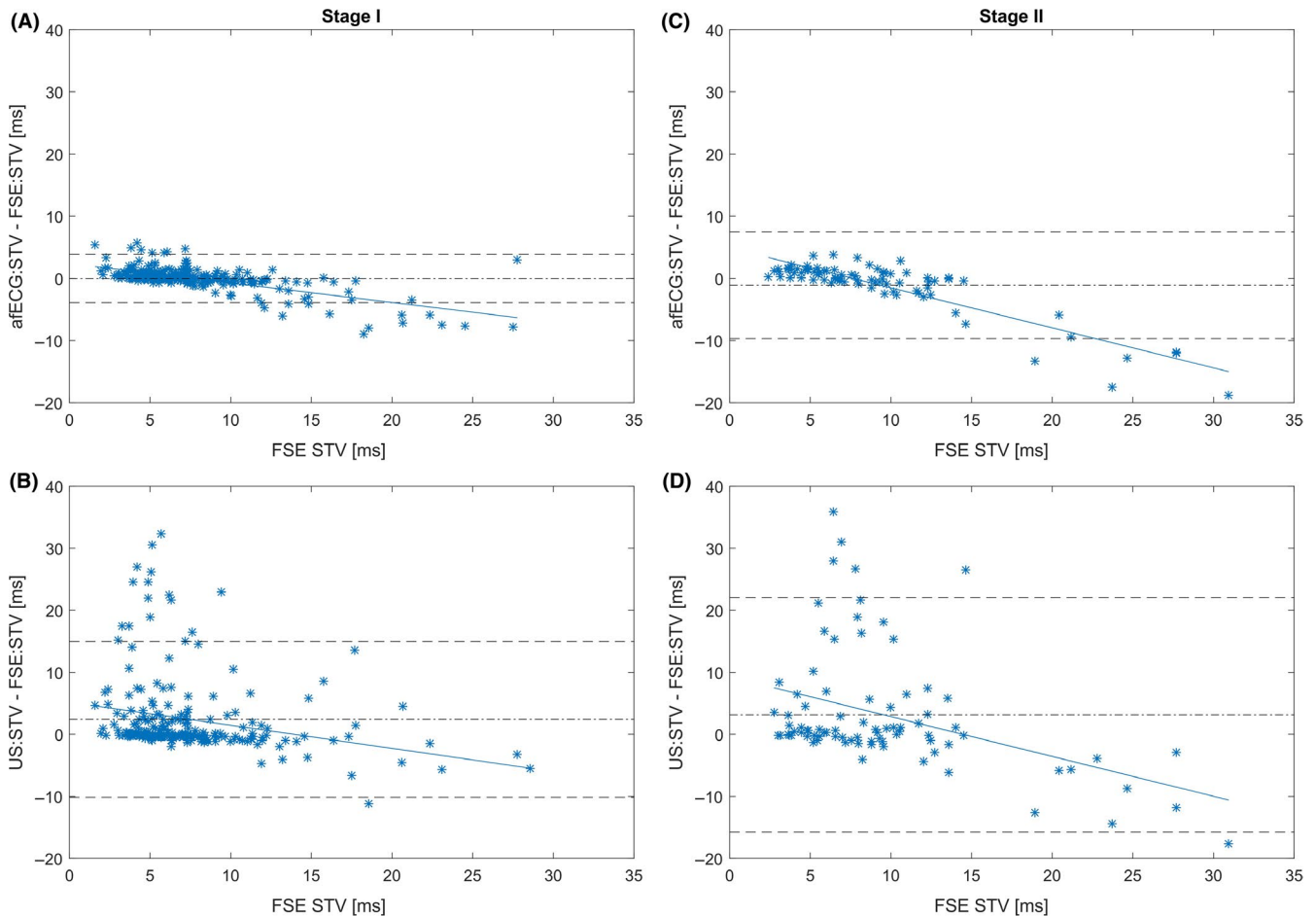
FIGURE 2 (A-D) Bland-Altman plots for short-term variability (STV) in stage I of labor comparing agreement of abdominal-fetal electrocardiogram (afECG) method with fetal scalp electrode (FSE) method (A) and Doppler ultrasound (US) with FSE (B). Stage II results are also shown for aFECG (C) and US (D). In both stages the afECG displayed better agreement with the FSE than did US [Color figure can be viewed at wileyonlinelibrary.com]

The results of all Bland-Altman analyses are summarized in Table 1. For each comparison, afECG gave results closer to those of the FSE than did US. Results of the comparison of the error distributions (Wilcoxon test) are shown in Table 2. For each parameter, with the exception of identifying large decelerations in the second stage, afECG was superior to US.

3.2 | Baseline rate

The spread of average baseline heart rate values around the nearly horizontal Bland-Altman regression line is quite narrow for afECG, as can be seen visually (Figure 1A) and confirmed by the minimal bias, narrow limits of agreement, and small root mean square error (Table 1). When, however, US was compared with the FSE the results were quite different (Figure 1B). The spread of values around the regression line was substantially greater, with wider limits of agreement, and greater bias and root mean square error. An error of more than 10 bpm occurred only 1.0% of the time with afECG; it occurred about 14% of the time with US. These differences were similarly distributed in stages I and II of labor (Figures 1C and 1D). For each comparison of errors (the difference between paired external and FSE-derived heart rates),

the differences between the afECG and US baseline FHR values were statistically significant (Table 2).

3.3 | Variability

For measures of both STV and LTV, the difference between the external method and the FSE-derived variability was substantially greater for the US than the afECG method, differences most evident at diminished variability levels, when error has the greatest clinical importance. This was true in both first and second stages of labor. When the STV from the FSE was low (<5.0 ms) the afECG error remained stable, but the US error increased considerably (Figures 2–3). The Bland-Altman LTV plots (Figures 3A–3D) were similar to those for STV in that the afECG error was minimal for low values of LTV.

3.4 | Accelerations and decelerations

The afECG method identified periodic small and large accelerations and small decelerations with greater reliability than did US
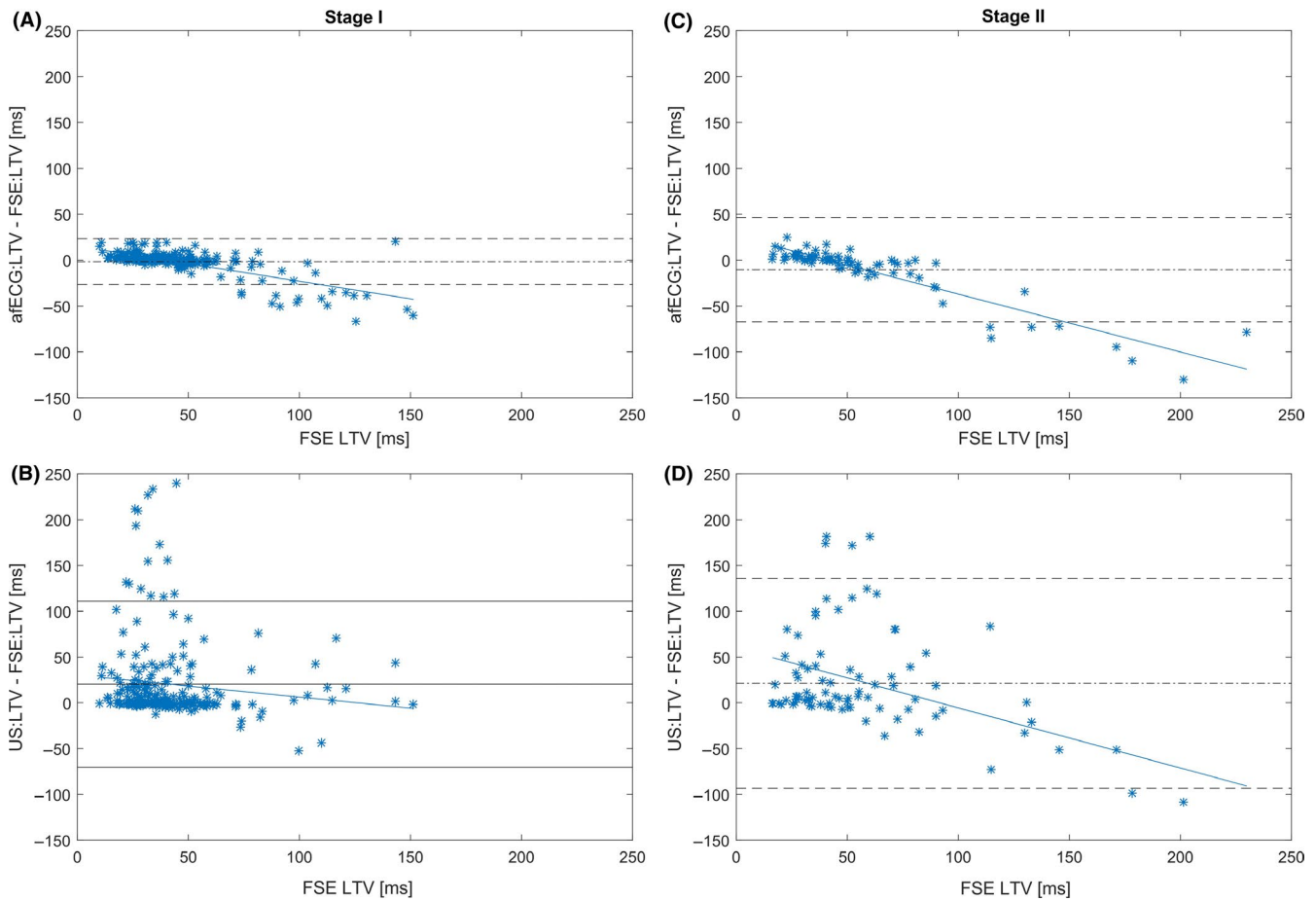
FIGURE 3 (A-D) Bland-Altman plots for long-term variability (LTV) in stage I of labor comparing agreement of abdominal-fetal electrocardiogram (afECG) method with fetal scalp electrode (FSE) method (A) and Doppler ultrasound (US) with FSE (B). Stage II results are also shown for aFECG (C) and US (D). In both stages the afECG displayed better agreement with the FSE than did US [Color figure can be viewed at wileyonlinelibrary.com]

in both stage I and stage II of labor. That is to say, the differences in the error frequencies between the external monitoring modes and the FSE were statistically significant, and Bland-Altman patterns of agreement showed a wider spread around the regression line for US than for afECG (Table 1). Large deceleration identifications in stage I showed that the errors between afECG and the FSE were smaller than those between US and the FSE ($P = 0.015$) (Table 2); however, for stage II they were not significantly different ($P = 0.255$).

Figures 4A–4D indicate the large deceleration error count as a function of frame number, demonstrating false positive identifications in Stages I and II for afECG and US. These results (Table 3) demonstrate overall a low rate of false events for both external monitoring methods, generally lower than 2 per 30 minutes. For both methods, the false-negative and false-positive rates were significantly higher in the second stage than in the first. afECG had a significantly higher rate of false-negative accelerations in both labor stages, and a higher false-negative rate for small decelerations in the second. False-positive accelerations and decelerations were significantly more common during the first stage with US than with afECG monitoring.

4 | DISCUSSION

We found the computerized identification of FHR characteristics for the afECG technique was in most instances more faithful to results from the scalp electrode than US-based FHR detection in the identification of baseline heart rate, and heart rate variability, accelerations, and decelerations. These results considerably expand our understanding of afECG-derived FHR patterns. Although the high accuracy and reliability of this technique in recognizing fetal cardiac activity was demonstrated previously, we have now shown that fundamental characteristics of heart rate patterns used in assessment of the fetus (rate, variability, periodic changes) are readily recognizable when using an afECG system. This system, therefore, proved equivalent or superior to US in replicating heart rate patterns derived from an FSE.

Our results have important clinical implications. Current recommendations for the clinical assessment of variability in the interpretation of FHR patterns require making fine distinctions in the degree of variability observed. Because such distinctions have a potentially substantial influence on decision-making, and because they are a component of all extant systems for FHR
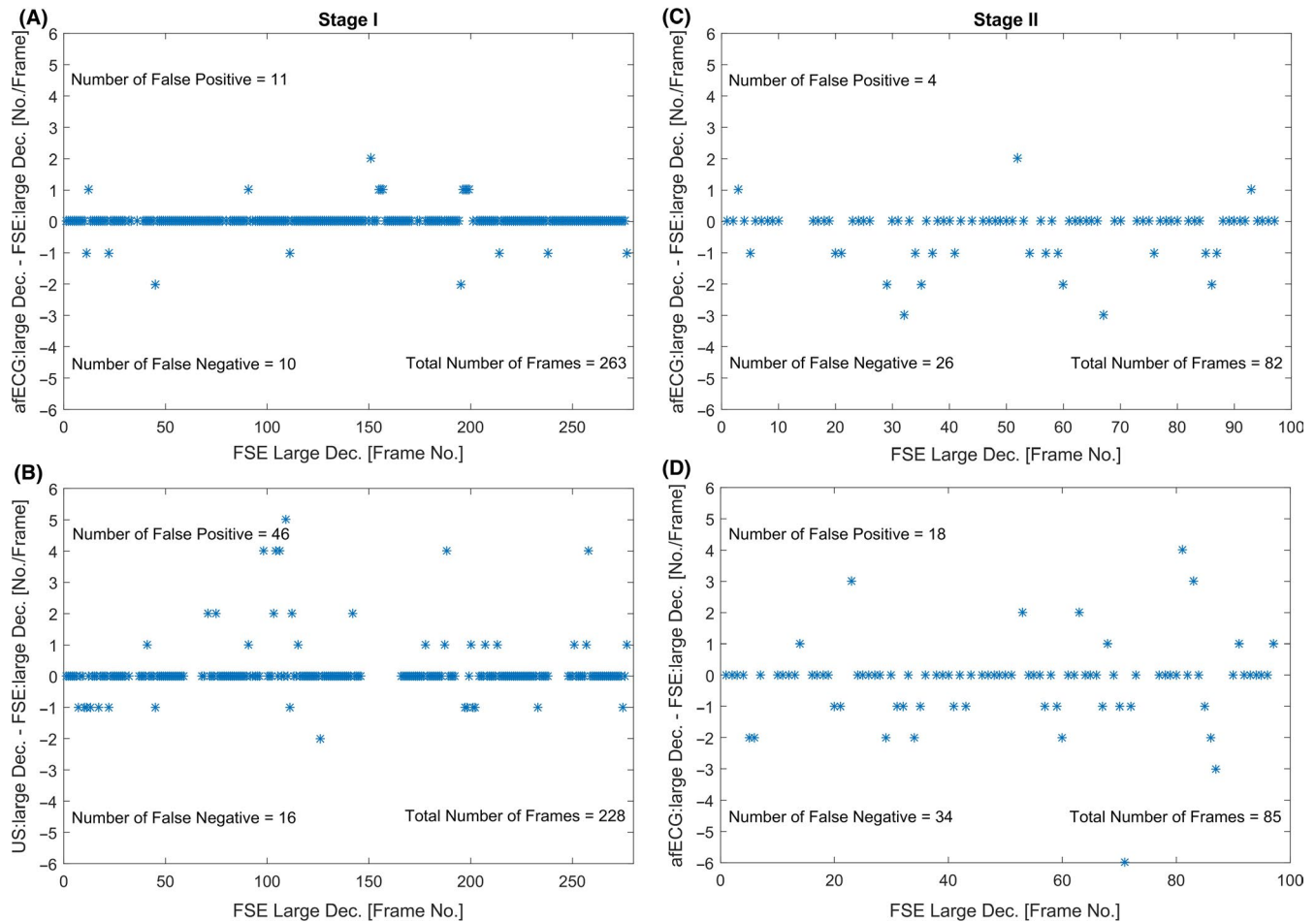
FIGURE 4 (A-D) Large deceleration error count as a function of frame number for abdominal-fetal electrocardiogram (afECG) (A and C) and ultrasound (US) (B and D) for both stages I and II. afECG shows consistently lower error count than US [Color figure can be viewed at wileyonlinelibrary.com]

pattern interpretation,[28-30] it is critically important that the technique used to assess rate and variability provide the most accurate available information. It has been assumed since the introduction of autocorrelation techniques to cardiotachometer software that US-derived FHR patterns can be assumed to display true FHR variability. Although this may be a reasonable generalization, it is not always correct.[31-34] On the contrary, our results show that the representation of FHR variability may be altered considerably by US-based monitoring, and should be interpreted with considerable caution. Moreover, the situations in which US is most likely to exaggerate true variability are those in which it is actually diminished, potentially leading to false reassurance and failure to timely intervention. If an external FHR monitoring technique is chosen, afECG will depict STV and LTV more accurately than will US techniques.

Our results also indicate that the clinician can be confident that periodic changes in the heart rate pattern—accelerations and decelerations of various descriptions—will be as recognizable with afECG-derived monitoring as with US. In fact, the error in identifying accelerations and decelerations was consistently smaller with afECG.

A large randomized trial would be necessary to determine whether use of this kind of monitoring technique results in neonatal outcomes that are equivalent or superior to monitoring with traditional US-derived heart rate patterns.

This is the first study to address whether afECG monitoring produces FHR patterns that would be interpreted in the same manner as those produced by US-based cardiotachometry. Of concern in designing this study was the fact that there is considerable inter- and intra-observer variation in the visual interpretation of FHR patterns.[35,36] We addressed this potential pitfall by using an established electronic means to identify FHR characteristics. This substituted an objective and completely uniform diagnostic tool for the inherent subjectivity of human visual interpretation.

This was a secondary analysis of a portion of data from a previously published study. The parent study was not designed or powered to address the specific hypothesis of this analysis. Nevertheless, we had a substantial number of observations, and almost all our results were statistically significant. This approach to FHR analysis does not distinguish among various types of decelerations, as visual observation can, and not every practitioner would agree with all features of the algorithm we used for FHR pattern identification. That notwithstanding, it provided us with a purely objective and unbiased

means of analyzing and comparing the outputs of the three modes of monitoring used in the study.

Another potential limitation was the fact that the US device was evaluated and adjusted (if necessary) every 20–30 minutes, and we did not document the number of adjustments made in each participant. This could have resulted in misdirected insonation and the potential for error in the inter-evaluation intervals, reinforcing the dangers of relying on US-derived measures of fetal well-being in a busy ward. This issue is not a concern with afECG monitoring, where the electrodes do not require adjustment during the entire labor, presenting the medical team with increased time to care for the mother rather than the electronic fetal monitoring instrument.

5 | CONCLUSION

We conclude that the afECG-based technique for continuous FHR monitoring is an appropriate choice for noninvasive monitoring. The heart rate patterns it generated were interpretable, and were generally superior to those produced by a traditional US method in terms of their conformity with patterns derived from a direct fetal electrode. The observed differences were generally of sufficient magnitude to influence the clinical interpretation of FHR patterns.

CONFLICTS OF INTEREST

Barrie Hayes-Gill and Terrence Martin were once Executive Director and Director of Marketing, respectively, at Monica Healthcare, Ltd. Neither has any current affiliation with the company or with GE Healthcare. Wayne R. Cohen has been a paid consultant to Monica Healthcare, Ltd. and to GE Healthcare.

REFERENCES

1. Reinhard J, Hayes-Gill BR, Schiermeier S, et al. Intrapartum signal quality with external fetal heart rate monitoring: a two-way trial of external Doppler CTG ultrasound and the abdominal fetal electrocardiogram. *Arch Gynecol Obstet*. 2012;286:1103-1107.
2. Georgieva A, Abry P, Chudáček V, et al. Computer-based intrapartum fetal monitoring and beyond: A review of the 2nd workshop on signal processing and monitoring in labor (October 2017, Oxford, UK). *Acta Obstet Gynecol Scand*. 2019;98:1207-1217.
3. Schramm K, Lapert F, Nees J, et al. Acceptance of a new non-invasive fetal monitoring system and attitude for telemedicine approaches in obstetrics: a case-control study. *Arch Gynecol Obstet*. 2018;298:1085-1093.
4. Sameni R, Clifford GD. A review of fetal ECG signal processing; issues and promising directions. *Open Pacing Electrophysiol Ther J*. 2010;1:4-20.
5. Cohen WR, Hayes-Gill B. Influence of maternal body mass index on accuracy and reliability of external fetal monitoring techniques. *Acta Obstet Gynecol Scand*. 2014;93:590-595.
6. Graatsma EM, Jacod BC, van Egmond LA, Mulder EJ, Visser GH. Fetal electrocardiography: feasibility of long-term fetal heart rate recordings. *BJOG*. 2009;116:334-337.
7. Jaros R, Martinek R, Kahankova R. Non-adaptive methods for fetal ECG signal processing: a review and appraisal. *Sensors*. 2018;18:3468-3482.
8. Cohen WR, Ommani S, Hassan S, et al. Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes. *Acta Obstet Gynecol Scand*. 2012;91:1306-1313.
9. Dawes GS, Moulden M, Redman CW. Criteria for the design of fetal heart rate analysis systems. *Int J Biomed Comput*. 1990;25:287-294.
10. Dawes GS, Moulden M, Redman CW. The advantages of computerized fetal heart rate analysis. *J Perinat Med*. 1991;19:39-45.
11. 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.
12. Dawes GS, Moulden M, Redman CWG. Shortterm fetal heart rate variation, decelerations, and umbilical flow velocity waveforms before labour. *Obstet Gynecol*. 1992;80:673-678.
13. Pardey J, Moulden M, Redman CWG. A computer system for the numerical analysis of nonstress tests. *Am J Obstet Gynecol*. 2002;186:1095-1103.
14. Serra V, Bellver J, Moulden M, Redman CWG. Computerized analysis of normal fetal heart rate pattern throughout gestation. *Ultrasound Obstet Gynecol*. 2009;34:74-79.
15. Pello LC, Rosevear SK, Dawes GS, Moulden M, Redman CW. Computerized fetal heart rate analysis in labor. *Obstet Gynecol*. 1991;78:602-610.
16. Dawes GS, Rosevear SK, Pello LC, Moulden M, Redman CW. Computerized analysis of episodic changes in fetal heart rate variation in early labor. *Am J Obstet Gynecol*. 1991;165:618-624.
17. Dawes GS, Moulden M, Redman CW. Improvements in computerized fetal heart rate analysis antepartum. *J Perinat Med*. 1996;24:25-36.
18. Nijhuis IJM, ten Hof J, Mulder EJH, et al. Numerical fetal heart rate analysis: nomograms, minimal duration of recording and interfetal consistency. *Prenatal Neonatal Med*. 1998;3:314-322.
19. Seliger G, Stenzel A, Kowalski EM, et al. Evaluation of standardized, computerized Dawes/Redman heart-rate analysis based on different recording methods and in relation to fetal beat-to-beat heart rate variability. *J Perinat Med*. 2016;44:785-792.
20. Kouskouti C, Jonas H, Regner K, Ruisinger P, Knabl J, Kainer F. Validation of a new algorithm for the short-term variation of the fetal heart rate: an antepartum prospective study. *J Perinat Med*. 2018;46:599-604.
21. Hiett AK, Devoe LD, Youssef A, Gardner P, Black M. A comparison of visual and automated methods of analysing fetal heart rate tests. *Am J Obstet Gynecol*. 1993;168:1517-1521.
22. Dawes GS, Moulden M, Redman CW. System 8000: computerized antenatal FHR analysis. *J Perinat Med*. 1991;19:47-51.
23. Redman CWG. *Sonicaid Fetalcare Antepartum Analysis Clinical Application Guide*. Cardiff, UK: Huntleigh Healthcare Ltd: 2003.
24. Dobbe JGG, Lunshof S, Boer K, Wolf H, Grimbergen CA. The technique and algorithms for computerized analysis of long-term fetal heart rate recordings. *Prenat Neonat Med*. 2001;6:280-289.
25. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *Statistician*. 1983;32:307-317.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
27. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)*. 2015;25:141-151.

28. Cohen WR, Yeh S-Y. The abnormal fetal heart rate baseline. *Clin Obstet Gynecol*. 1986;29:73-82.

29. Warmerdam GJJ, Vullings R, Van Laar J, et al. Detection rate of fetal distress using contraction-dependent fetal heart rate variability analysis. *Physiol Meas*. 2018;39:025008.

30. Garabedian C, Butruille L, Drumez E, et al. Inter-observer reliability of 4 heart rate classification systems. *J Gynecol Obstet Hum Reprod*. 2017;46:131-135.

31. Fukushima T, Flores CA, Hon EH, Davidson EC Jr. Limitations of autocorrelation in fetal heart rate monitoring. *Am J Obstet Gynecol*. 1985;153:685-692.

32. Divon MY, Torres FP, Yeh SY, Paul RH. Autocorrelation techniques in fetal monitoring. *Am J Obstet Gynecol*. 1985;151:2-6.

33. Nunes I, Ayres-de-Campos D, Costa-Santos C, Bernardes J. Differences between external and internal fetal heart rate monitoring during the second stage of labor: a prospective observational study. *J Perinat Med*. 2014;42:493-498.

34. Goncalves H, Rocha HP, Ayres-de-Campos D, Bernardes J. Internal vs. external intrapartum foetal heart rate monitoring: the effect on linear and nonlinear parameters. *Physiol Meas*. 2006;27:307-319.

35. Figueras F, Albela S, Bonino S, et al. Visual analysis of antepartum fetal heart rate tracings: inter- and intra-observer agreement and impact of knowledge of neonatal outcome. *J Perinat Med*. 2005;33:241-245.

36. Blackwell SC, Grobman WA, Antoniewicz L, Hutchinson M, Bannerman CG. Interobserver and intraobserver reliability of the NICHD 3-tier fetal heart rate interpretation system. *Am J Obstet Gynecol*. 2011;205:378.e1-5.