

1 **Combination of XGBoost Analysis and Rule-Based**
2 **Method for Intrapartum Cardiotocograph Classification**

3
4 Pao-Lin Kuo¹, Lim Bee Yen², Yi-Chun Du^{2, *}, Po-Fan Chen¹, Pei-Yin Tsai¹

5
6 ¹ Department of Obstetrics and Gynecology, College of Medicine, National Cheng Kung
7 University, Tainan 70105, Taiwan

8 ² Department of Biomedical Engineering, National Cheng Kung University, Tainan 70105,
9 Taiwan; Medical Device Innovation Center, National Cheng Kung University, Tainan 70105,
10 Taiwan

11
12
13
14
15
16
17
18
19 *Corresponding author: Yi-Chun Du, Ph.D.

20 Email: terrydu@gs.ncku.edu.tw

23 **Abstract:**

24

25 Cardiotocograph (CTG) contains uterine contraction (UC) and fetal heart rate (FHR)
26 signals, which is an important information of clinical pregnant woman care. National Institute
27 of Child Health and Human Development (NICHD) is one of the reference guides for clinical
28 care and it classified pregnant woman into three categories including I, II, and III to evaluate
29 the status of fetus. However, when it comes to manual interpretation was time-consuming and
30 not easy to observe the slight differences. In this study, we combined rule-based method and
31 eXtreme Gradient Boosting (XGBoost) analysis for intrapartum cardiotocograph
32 classification. Because the category II of NICHD is defined unknown status, XGBoost
33 analysis was used to classify the category II into IIa and IIb, and analyze their probability of
34 fetal distress (FD). From the clinical trial of 68 pregnant women, the results of three
35 categories (I, II and III) were consistent and no statistical difference with the clinicians'
36 interpretation and the average Kappa was about 0.72. The results also indicated that the
37 probability of FD was 28.8% and 71.2% in category IIa and IIb, respectively. It shows the
38 proposed method has potential to provide a clinical assistant tool for pregnant women care.

39

40 **Keywords:** Cardiotocograph (CTG), National Institute of Child Health and Human
41 Development (NICHD), eXtreme Gradient Boosting (XGBoost), Fetal Distress (FD)

42

43

44 1. Introduction

45 Advanced maternal age (AMA) is defined as a phenomenon that women become
46 pregnant at the age above 35, and it has increased in high-income economy countries recently
47 [1]. In 2018, a research from Pew Research Center in America revealed that 86% of American
48 women had ever given birth at ages 40 to 44 [2]. A study by EU researchers indicates that
49 20% of birth in England and Wales were given by women ages above 35 years, while 4%
50 were given by women ages ≥ 40 years, compared with that of 6% and 1% birth respectively in
51 1980 [1]. The most general reason of this trend is attributed to the advancements in assisted
52 reproductive technologies (ART). However, this trend potentially causes clinic risks [1].
53 According to previous studies [3], advanced maternal age is associated with the incidence of
54 complications of pregnancy, such as abortion, gestational toxemia, gestational diabetes, and
55 fetal growth retardation. These complications not only trouble pregnant women but also affect
56 the health of the fetuses. Among all these complications, fetal distress (FD) may lead to
57 hypoxia of the fetus, thereby reducing fetal movements, reducing fetal heart rate, and
58 provoking acidemia. In addition, Hypoxic Ischemic Encephalopathy (HIE) caused by FD may
59 also lead to chronic, long-term injury or even death to the fetus [4]. The placenta is the main
60 source of nutrition for the fetus, while the function of the uterus gradually declines with the
61 age of women [5]. In 2014, scholar Dr. Parker, S. pointed out that advanced maternal age
62 would cause an increasing chance of early placental exfoliation, which would reduce the
63 oxygen circulation rate of fetus and thus provoke FD [6] and may lead to growth retardation.
64 According to statistics, the incidence of FD during delivery is three out of a thousand [7],
65 therefore, in 2013, Dr. Austin Ugwumadu et al. looked for the relevant fetal compensation
66 strategies through observation of the changes in fetal heart rate during pregnancy, hoping to
67 lower the incidence of FD [8].

68 Other studies have shown that there is a correlation between prenatal and fetal
69 physiological signals and FD [9-15], for example, cardiotocograph (CTG) signals can assist
70 obstetricians and nurses in clinical interpretation. CTG is mainly used as a clinical reference
71 for fetal health and FD [9] by monitoring the changes of Uterine Contraction (UC) and Fetal
72 Heart Rate (FHR) of pregnant women. However, in recent years many experts and scholars or
73 organizations, for example, National Institute of Child Health and Human Development
74 (NICHD) [10] and Royal College of Obstetricians and Gynecologists (RCOG) [11], have
75 established different classification rules based on different clinical experiences in using CTG
76 signals, so as to make clinical results more objective. In order to improve the speed of
77 interpretation and decrease human error, various methods of intelligent CTG interpretation

78 have been proposed in recent years. According to the definition of NICHD, the research team
79 from National Taiwan University carried out the development of automated NICHD 3-tier
80 classification system in 2014. In this study, researchers collected 62 CTG data before delivery
81 and compared the CTG data with the visual interpretations from 8 obstetricians, and the
82 results showed the average statistics kappa value of category I, II, III, and overall were
83 respectively 0.89, 0.78, 0.50 and 0.79. Compared with the traditional classification method,
84 the kappa value here was higher (0.15-0.38), but the value of category III was lower [12]. In
85 2016, according to RCOG, a Malaysian research team classified a group of pregnant women
86 into normal, suspicious and pathological categories, and analyzed the baseline, variability,
87 deceleration, acceleration based on FHR to construct an instant classification system [11]. The
88 proposed methods were employed to confirm the accuracy of classification based on the
89 visual interpretation of three obstetricians, kappa values of 0.668, 0.587 and 0.630 for the 80
90 clinical CTG signals.

91 As it can be seen from the above, the automatic interpretation system of NICHD or
92 RCOG has its advantages and feasibility; however, some classification results are clinically
93 vague, such as the category II of NICHD and the suspicious category of RCOG. In 2016,
94 Penfield, CA et al. proposed a classification system for NICHD category II signals, called
95 “ABC system,” which divided category II tracings into IIA, IIB, and IIC subcategories. The
96 system mainly classified the changes in the new Fetal heart rate tracing characteristics in
97 category II. They demonstrated that the ABC system could improve the team communication
98 and increase on-site management of category II FHR tracings by private physicians [13]. In
99 2017, Sabina Martı Gamboa et al. started to estimate the association between atypical variable
100 decelerations and neonatal acidemia and compare the CTG signals of 102 cases of fetal
101 acidemia (umbilical arterial cord gas pH \leq 7.10) and of 100 normal fetuses (umbilical arterial
102 cord gas pH $>$ 7.10), so as to estimate the correlation among typical variable decelerations of
103 CTG, neonatal acidemia, and the morbidity of FD [14]. The results show that certain atypical
104 features, as slow return and loss of moderate variability within decelerations are associated
105 with neonatal acidemia. The slow return could help in the gradation of acidemia risk levels, as
106 an indicator of gravity.

107 In view of this, this study used the classification rules of NICHD with rule-based method
108 to classify our results into category I, II, and III, and by literature results [14], the atypical
109 variable decelerations directly into category III, e.g. slow-return, to improve its consistency
110 with the interpretation by clinicians. In addition, every UC during pregnancy is a kind of
111 stimulation to the fetus; for normal fetuses, such stimulation can be recovered in a short time,
112 however, if the fetus has FD, the fetus may not be able to maintain internal regulation through

113 compensation strategies [15].

114 Extreme Gradient Boosting (XGBoost) is an open-source library that provides an
115 efficient and effective implementation of the gradient boosting analysis. XGBoost is a class of
116 ensemble machine learning algorithms that can be used for classification predictive modeling
117 problems. In this study, we utilized XGBoost analysis as the classifier based on sequence
118 analysis to classify category II data into IIa and IIb by further examining UC and FHR. It
119 could improve the understanding of category II signals in clinical application, increase on-site
120 management of category II by private physicians, and provide a more accurate clinical basis
121 for FD.

122 **2. Materials and Methods**

123 The proposed system for intrapartum CTG classification was developed by rule-based
124 method and XGBoost analysis in this study. Data sources were mainly obtained from a
125 clinical database and then analyzed accordingly. The clinical CTG signals were divided into
126 category I, II & III by the rule-based method on the features such as baseline, variability,
127 acceleration, and deceleration following NICHD and Chen, C. Y. [12] definition. In order to
128 reduce the processing time, we modified and simplices the rule as Table 1.

129 **2.1. Clinical Data Collection**

130 The clinical CTG records were taken from the fetal monitor (Avalon FM20, Royal
131 Philips, Eindhoven, Netherlands) in the delivery room of the National Cheng Kung University
132 Hospital. The medical records were documented between January 2017 and April 2018. The
133 excluding rules of clinical CTG records were Sinusoidal pattern of FHR (SP), multiple births,
134 the mother has heart and lung diseases, and the fetus has congenital anomalies. The flowchart
135 of clinical CTG selection is shown in Figure 1. The data, which included numerical data in the
136 database and converted images, comprised of the age of the pregnant woman, gestational age,
137 FHR, UC, birthing methods, and anesthesia methods, as shown in Table 2. The sixty-eight
138 records of UC and FHR were selected in this study and the data included 30 & 30 records were
139 with FD & Non-FD, respectively, that were verified by professional medical staffs. The
140 information was obtained under the specification of the Institutional Review Board (IRB)
141 program (A-ER-105-477). Each record was continuous measurement at least 40 minutes with
142 1 Hz sampling rate.

143 **2.2. The Preprocessing of CTG Signals**

144 During prenatal checkups, CTG records were affected easily by the noise and missing
145 data. It may cause by the movement of pregnant woman or sensors which greatly affects the

146 identification results made by the medical staff or computerized analysis [11]. Therefore,
 147 before performing feature extraction, we referred the research [16-17] for CTG preprocessing
 148 that included linear interpolation, trend removes and data smoothing to remove the noise and
 149 motional artifact.

150 **2.3. Features Extraction of FHR and UC**

151 Features extraction from clinical CTG data was import in this study following Table 1.
 152 At first, various types of features were extracted from the FHR signal like baseline, variability,
 153 acceleration, and deceleration for instance, the time difference between FHR & UC as a
 154 feature. Five parameters were defined from the extraction of the time and frequency domains,
 155 as well as the morphological features. In this section, FHR features were extracted based on
 156 the NICHD criteria [12]. The algorithm used for feature extraction and classification in this
 157 study was developed and program implemented by MATLAB (version 9.9.0.1495850,
 158 R2020b, Natick, Massachusetts, The MathWorks Inc.).

159 A. Baseline (BL) of FHR detection:

160 According to the definition of NICHD, the BL of FHR is determined by approximating
 161 the mean FHR rounded to increments of 5 beats per minute (bpm) during a 10-minute window.
 162 Moreover, it excluding the state of accelerations and decelerations. There must be at least 2
 163 minutes of identifiable baseline segments (not necessarily contiguous) in any 10-minute
 164 window, normal (BLN), tachycardia (BLT) and bradycardia (BLB) were also determined, as
 165 shown in Equation (1). Where the BV in the above formula represents the difference between
 166 the maximum and minimum values of the FHR within one minute; i represents the FHR
 167 reference line in the i -th minute; and n is the total number of times the difference before and
 168 after the FHR is less than 5 bpm.

$$\begin{aligned}
 169 \quad & \text{if } BV \leq 5, BL_{(i)} = \frac{1}{n} \sum_1^n FHR_{(n)} \\
 170 \quad & \text{if } BV > 5, BL_{(i)} = BL_{(i-1)} \\
 171 \quad & BV = FHR_{max}^{(n)} - FHR_{min}^{(n)} \tag{1}
 \end{aligned}$$

172 B. Baseline variability (BV) of FHR detection:

173 The FHR BV is a response that shows FHR changes within a short period of time when
 174 the fetal central nervous system is stimulated. It also reflects the normal development of fetal
 175 nerves and the adequacy of oxygen in the body [15]. Through the BV, we could observe
 176 whether the abnormal heart rate of the fetus was due to metabolic acidosis or neurological
 177 damage that maybe caused by hypoxia. FHR variability with 0 beats per minute (bpm) does
 178 not exist in the living fetus, and we used the definition of Chen et al., with an amplitude of 2
 179 bpm as BVA, an amplitude between 2-5 bpm as BVMI, it means that the fetus may have been

180 neurologically harmed due to hypoxia [15]. An amplitude between 6-25 bpm as BVMO, an
 181 amplitude more than 25 bpm as BVMA [12]. The BV is calculated by the equation (2) of the
 182 difference between the FHR maximum and minimum values in every minute. n represents one
 183 minute. Moreover, the BV would exclude the AC and DE condition.

$$184 \quad BV = FHR_{max}^{(n)} - FHR_{min}^{(n)} \quad (2)$$

185 C. Acceleration (AC) of FHR detection:

186 The manifestation of the FHR acceleration is closely related to fetal health and is usually
 187 generated when there is fetus activity. We compare the lowest point of acceleration with the
 188 peak position of UC and calculate the continuous acceleration time of FHR. The duration and
 189 magnitude of the acceleration during measurement are used as evaluation criteria. Generally,
 190 the duration is more than 15 seconds and the amplitude are greater than 15 bpm, as shown in
 191 Equation (3). FHR onset & FHR recovery means the start and the end of continuous
 192 acceleration time, respectively. The FHR acceleration represents the normal oxygen
 193 concentration in the blood of the fetus with no metabolic academia; therefore, the fetus is
 194 currently in a normal state. However, during the monitoring process, when no acceleration is
 195 observed, it does not mean that the fetus is abnormal. The determination must be done
 196 through FHR variability and deceleration.

$$197 \quad FHR_{max} - BL \geq 15 \text{ bpm}$$

$$198 \quad \text{and } FHR_{recovery} - FHR_{onset} \geq 15 \text{ sec}$$

199 (3)

200 D. Deceleration (DE) of FHR detection:

201 UC is the state in which the uterus practices childbirth during pregnancy. When the
 202 uterus contracts, it is squeezed inwardly, and the blood flow in the placenta changes. We used
 203 the valley detection method is based on an algorithm that fits a quadruple polynomial to
 204 sequential groups of data points, to determine the nadirs of FHR decelerations [13]. At this
 205 moment, the DE is the degree of FHR decline that is consistent with the UC. The DE standard
 206 is longer than 15 seconds and the deceleration size is more than 15 bpm, as shown in Equation
 207 (4). The nadir location of FHR deceleration was compared with the UC peak location to
 208 classify early and late decelerations.

$$209 \quad BL - FHR_{min} \geq 15 \text{ bpm},$$

$$210 \quad \text{and } FHR_{recovery} - FHR_{onset} \geq 15 \text{ sec}$$

211 (4)

212 E. Early DE (EDE) & Late DE (LDE) of FHR detection:

213 When the uterus contracts under normal conditions, the fetus initiates a compensatory
 214 mechanism to balance the oxygen in the body. Generally, the time difference between the

215 lowest point of FHR and the last max value before the lowest point of FHR lasts at least 30
 216 seconds, and there is no time difference between the position of FHR_{nadir} , which is the lowest
 217 position of FHR deceleration, and the position of UC_{max} . After the end of the contraction, the
 218 FHR returns to normal. Such stimulation does not affect the fetus and belongs to the state of
 219 EDE, as shown in Equation (5).

220 The state of late deceleration reflects the condition where the fetus may face hypoxia.
 221 Therefore, the FHR is delayed and decelerated after UC. LDE is defined as that FHR starts to
 222 decelerate to the bottom at least 30 seconds, and its time difference from UC starts is more
 223 than or equal to 18 seconds, as shown in Equation (6).

$$\begin{aligned}
 224 \quad EDE &= FHR_{onset}(n) - FHR_{nadir}(n) \geq 30 \text{ sec} \\
 225 \quad \text{and } FHR_{nadir}(n) &= UC_{max}(n) \\
 226 \quad LDE &= FHR_{onset}(n) - FHR_{nadir}(n) \geq 30 \text{ sec} \\
 227 \quad \text{and } UC_{onset}(n) - FHR_{onset}(n) &\geq 18 \text{ sec} \\
 228 \quad (6)
 \end{aligned}
 \tag{5}$$

229 F. UC Duration:

230 UC feature samples are firstly obtained by setting the threshold value (Thr). When the
 231 threshold of the UC data is extracted, and the UC is greater than the threshold value, an
 232 intersection point is generated. The time interval between the two intersection points (a, b) is
 233 used as the basis for judging the UC, as shown in Equation (7). The difference between the
 234 time point at which the maximum occurs and the time when the FHR decelerates can be used
 235 to determine whether the FHR deceleration state is early deceleration or late deceleration.

$$\begin{aligned}
 236 \quad UC_{var} &= UC_{max}(n) - UC_{min}(n) \\
 237 \quad \text{if } UC_{var} \leq 5, \quad UC_{Thr}(i) &= \frac{1}{n} \sum_{i=1}^n UC_{(i)} \\
 238 \quad \text{else } UC_{var} > 5, \quad UC_{Thr}(i) &= UC_{Thr}(i - 1) \\
 239 \quad \text{and } t_{a \rightarrow b} \geq 30, \quad UC &= UC_{max}(t_{a \rightarrow b}) \\
 &
 \end{aligned}
 \tag{7}$$

240 **2.4. Signal Truncation for Sequence Analysis and XGBoost Analysis**

241 In this study, we design a CTG category II advanced analysis method based on UC to
 242 further classify the signals into IIa and IIb. Since the traditional method requires a long-term
 243 evaluation such as EDE & LDE in the interpretation of CTG, it is not applicable in the
 244 interval of a UC, and therefore only the parameters of BL, BV, AC, and DE in each interval
 245 are taken in this study. In the sequence analysis, based on UC, the CTG signal is evaluated as
 246 normal and abnormal FHR through truncation and the above parameters, and the ratio of the 2
 247 CTG signals is used to further classify the category II signal, shown in Figure 2. In this study,
 248 we applied a machine learning classifier, the XGBoost [18], to classify the category II CTG
 249 signals following NICHD rules and further classify into IIa and IIb. The input features

250 included 4 characteristics of FHR and UC obtained by the rule-based method (BL, BV, AC,
251 DE), and the abnormal ratio in the sequence analysis. In addition, we take the results from
252 clinicians as the ground truth of the classifier training model. This method was chosen due to
253 its significant advantages, including dealing with missing values, requiring data scaling and
254 implying a computationally efficient variant of gradient boosting analysis [18]. It provided
255 satisfactory results in ML competitions and was successfully used in other studies and
256 domains [19, 20]. We can efficiently select features parameters, reduce overfitting and the
257 computational complexity with supervised learning of XGBoost. Through supervised learning
258 and additive training to execute iteration and update the objective function, then build a model
259 after training to predict the results. Objective function of XGBoost as shown in Equation (8)
260 and (9)

$$261 \quad \text{Obj}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (8)$$

$$262 \quad \text{where } \Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (9)$$

263 where l is loss function, to estimate the difference between \hat{y}_i and y_i , which is the training
264 error. Ω is regularization term, including T and γ denote decision node and hyperparameters,
265 respectively. By adjusting ω , the weight of decision node, to avoid overfitting.

266 3. Results

267 The data collected in this study was mainly CTG records provided by the clinical
268 database under IRB program. There were two kinds of CTG data, FD and Non-FD, that were
269 accordance with the clinical medical records. After backtracking the CTG information, the
270 data was given to three clinicians, their average obstetric experience more than 10 years, to
271 interpret the relevant CTG records and classify into category I, II and III that followed
272 NICHD. Then, we utilized the rule-based method to classify the data and calculated the ratio
273 of FD in each category. The results made by the clinicians and rule-based method were
274 compared to evaluate the accuracy. The comparison results of Kappa in each category were
275 shown in Table 3, and the average Kappa was about 0.72. There was no statistical difference
276 between the results from the clinicians and rule-based method.

277 In order to improve the clinical evaluation of categories II of NICHD, we tried to classify
278 the category II into IIa and IIb and to analyze their correlation of the probability of FD.
279 Category IIa means that the fetus has a higher probability of returning to the category I state,
280 and IIb means that the fetus has a higher probability of developing FD, thus, surgery must be
281 performed immediately. This processing method not only utilized the results of the sequence
282 analysis but also included the features of the rule-based extraction, which was used as an
283 input in the XGBoost analysis. The sequence analysis mainly used the UC time as the

284 standard when extracting the FHR segment. The FD ratio was calculated through Equation
285 (10). The results from the XGBoost analysis and classification were shown in Table 4. It can
286 be found that the FD ratio of IIa is significantly lower than IIb, and there is a statistical
287 difference ($p < 0.05$). Although I and IIa, IIb and III are not statistically consistent, clinicians
288 still can take IIa and IIb as the basis for clinical care.

$$289 \quad \text{FD ratio} = \frac{\text{Number of FD}}{\text{Quantity in the Category}}$$

290 (10)

291 **4. Discussion**

292 Based on clinical experience, clinicians diagnose the fetal status such as FD by CTG
293 changes with visual analysis, usually dominated near the end of CTG signal period. It is not a
294 very objective method if without consideration of all the long-term CTG signal (more than 40
295 minutes) [4]. The results of CTG interpreted by clinicians were compared with the proposed
296 system, we found that there were statistically consistent with the clinicians' visual analysis
297 through the Kappa value. However, we further analyze the different signals that compared
298 between clinicians and proposed system. We found out the difference usually caused by the
299 local and global view of CTG signal. For example, there was a result put under category I by
300 clinicians' visual analysis; nevertheless, in global view, once its variability met LDE, the
301 proposed system did not classify it as category I but category II. Since the long-term CTG
302 signal consideration that more than 40 mins by visual analysis, it is time-consuming and
303 difficult to execute in the clinical environment. A relative objective comment based on
304 long-term CTG analysis by the proposed system was necessary for clinical evaluation. The
305 categories of IIa and IIb was carried out by the XGBoost analysis improved the information of
306 fetus in original category II. The FD ratio from original category II by rule-based method was
307 about 65.2%, and after further classify by XGBoost analysis into categories IIa and IIb, the
308 FD ratio was changed to 28.8% and 71.2%, respectively. It could provide much reliable
309 information for clinical evaluation. The sequence analysis in this study not only contained the
310 original part of the feature analysis in rule-based method, but also gained a new feature by the
311 abnormal ratio of all the CTG records. However, the definition of variable DE is not clear at
312 present, it is also difficult to define a good feature extraction method in our system. Therefore,
313 more clinical trials and big-data set for further analysis and verification are required.

314 **5. Conclusions**

315 Combine the rule-based method and XGBoost analysis, it could improve clinical
316 evaluation of the CTG signal. According to the literature [21], in some cases of variable DE
317 such as slow return that was highly correlated with FD, it was also correlated with the

318 magnitude and the shape of the *pH* value of the fetal cord blood. This was significantly
319 correlated with fetal acidemia statistically. In future studies, the proposed method can be
320 combined with other physiological parameters of the fetus for further clinical evaluation to
321 improve the interpretation. Overall, the proposed method has highly potential to be an
322 assistive and warning system for pregnant women care to improve the care quality and reduce
323 the burden of medical staff.

324 **Acknowledgements:** The authors are very grateful for the assistance and co-operation of all
325 the patients and medical staff involved in the clinical test.

326 **Funding:** This study was supported by the Ministry of Science and Technology, Taiwan,
327 under grant MOST 107-2221-E-006 -240 -MY3 and MOST 108-2221-E-006 -231 -MY3.

328

329 **Compliance with Ethical Standards**

330

331 **Conflicts of Interest:** The authors declare that they have no known competing financial
332 interests or personal relationships that could have appeared to influence the work reported in
333 this paper.

334 **Ethical Approval:** The experimental protocol was established, according to the ethical
335 guidelines of the Helsinki Declaration and was approved by the Human Ethics Committee of
336 National Cheng Kung University Hospital (IRB Number: A-ER-105-477).

337 **Informed Consent:** Written informed consent was obtained from individual or guardian
338 participants.

339

340

341

342 References

- 343 [1] Lean, S. C., Derricott, H., Jones, R. L., & Heazell, A. E. (2017). Advanced maternal age
344 and adverse pregnancy outcomes: A systematic review and meta-analysis. *PloS one*,
345 12(10), e0186287.
- 346 [2] [http://www.pewsocialtrends.org/2018/01/18/theyre-waiting-longer-but-u-s-women-](http://www.pewsocialtrends.org/2018/01/18/theyre-waiting-longer-but-u-s-women-today-more-likely-to-have-children-than-a-decade-ago/)
347 today-more-likely-to-have-children-than-a-decade-ago/
- 348 [3] Khalil, A., Syngelaki, A., Maiz, N., Zinevich, Y., & Nicolaides, K. H. (2013). Maternal
349 age and adverse pregnancy outcome: a cohort study. *Ultrasound in Obstetrics &*
350 *Gynecology*, 42(6), 634-643.
- 351 [4] Martin CB, Jr. Normal fetal physiology and behavior, and adaptive responses with
352 hypoxemia. *Semin Perinatol* 2008; 32:239-42.
- 353 [5] Fowden, A. L., Coan, P. M., Angiolini, E., Burton, G. J., & Constancia, M. (2011).
354 Imprinted genes and the epigenetic regulation of placental phenotype. *Progress in*
355 *biophysics and molecular biology*, 106(1), 281-288.
- 356 [6] Parker, S. E., & Werler, M. M. (2014, April). Epidemiology of ischemic placental
357 disease: a focus on preterm gestations. In *Seminars in perinatology* (Vol. 38, No. 3, pp.
358 133-138). Elsevier.
- 359 [7] E. M. Graham, K. A. Ruis, A. L. Hartman, F. J. Northington, and H. E. Fox, “A
360 systematic review of the role of intrapartum hypoxia-ischemia in the causation of
361 neonatal encephalopathy,” *American Journal of Obstetrics and Gynecology*, vol. 199,
362 no. 6, pp. 587–595, 2008.
- 363 [8] Ugwumadu, A. (2013). Understanding cardiocographic patterns associated with
364 intrapartum fetal hypoxia and neurologic injury. *Best Practice & Research Clinical*
365 *Obstetrics & Gynaecology*, 27(4), 509-536.
- 366 [9] Alfirevic, Z., Devane, D., Gyte, G. M., & Cuthbert, A. (2017). Continuous
367 cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal
368 assessment during labour. *The Cochrane Library*.
- 369 [10] Macones, G. A., Hankins, G. D., Spong, C. Y., Hauth, J., & Moore, T. (2008). The 2008
370 National Institute of Child Health and Human Development workshop report on
371 electronic fetal monitoring: update on definitions, interpretation, and research guidelines.
372 *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 37(5), 510-515.
- 373 [11] Alyousif, S., Mohd, M. A., Bilal, B., Sheikh, M., & Algunaidi, M. (2016). Rule-Based
374 Algorithm for Intrapartum Cardiotocograph Pattern Features Extraction and
375 Classification. *Health Science Journal*, 10(6), 1.

- 376 [12] Chen, C. Y., Yu, C., Chang, C. C., & Lin, C. W. (2014). Comparison of a novel
377 computerized analysis program and visual interpretation of cardiotocography. *PloS one*,
378 9(12), e112296.
- 379 [13] Penfield, C. A., Hong, C., Ibrahim, S. E. H., Kilpatrick, S. J., & Gregory, K. D. (2016).
380 a system to stratify category II fetal heart rate tracings. *American journal of*
381 *perinatology*, 33(07), 688-695.
- 382 [14] Martí Gamboa, S., Pascual Mancho, J., Lapresta Moros, M., Rodríguez Lázaro, L., de
383 Bonrostro Torralba, C., & Castán Mateo, S. (2017). Atypical decelerations: do they
384 matter? *The Journal of Maternal-Fetal & Neonatal Medicine*, 30(2), 240-244.
- 385 [15] Williams Obstetrics, 24th Edition
- 386 [16] Wrobel, J., Roj, D., Jezewski, J., Horoba, K., Kupka, T., & Jezewski, M. (2015).
387 Evaluation of the robustness of fetal heart rate variability measures to low signal quality.
388 *Journal of Medical Imaging and Health Informatics*, 5(6), 1311-1318.
- 389 [17] Romano, M., Bifulco, P., Ponsiglione, A. M., Gargiulo, G. D., Amato, F., & Cesarelli,
390 M. (2018). Evaluation of floatingline and foetal heart rate variability. *Biomedical Signal*
391 *Processing and Control*, 39, 185-196.
- 392 [18] Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. In: Proceedings of
393 the 22Nd ACM SIGKDD international conference on knowledge discovery and data
394 mining. ACM, pp 785–794
- 395 [19] Torlay, L., Perrone-Bertolotti, M., Thomas, E., & Baciú, M. (2017). Machine
396 learning–XGBoost analysis of language networks to classify patients with epilepsy.
397 *Brain informatics*, 4(3), 159-169.
- 398 [20] Ogunleye, A. A., & Qing-Guo, W. (2019). XGBoost model for chronic kidney disease
399 diagnosis. *IEEE/ACM transactions on computational biology and bioinformatics*.
- 400 [21] Hamilton, E., Warrick, P., & O’Keeffe, D. (2012). Variable decelerations: do size and
401 shape matter. *The Journal of Maternal-Fetal & Neonatal Medicine*, 25(6), 648-653.
- 402

Table 1. The rules definition of categories I, II & III based on NICHD

CTG	FHR & UC
Signal preprocess	Zero Deletion
	Denoising
	Signal segmentation
Features extration	Baseline of FHR (BL): Normal (BLN), $110 \geq BL \geq 160 \text{ bpm} \ \& \ \geq 10 \text{ min}$ Tachycardia (BLT), $BL > 160 \text{ bpm} \ \& \ \geq 10 \text{ min}$ Bradycardia (BLB), $BL < 110 \text{ bpm} \ \& \ \geq 10 \text{ min}$
	Baseline variability of FHR (BV): unit = bpm Absent (BVA), $BV < 2$ Minimal (BVMI), $5 \geq BV \geq 2$ Moderate (BVMO), $25 \geq BV > 6$ Marked (BVMA), $BV > 25$
	Acceleration of FHR (AC): Onset to nadir $\geq 15 \text{ sec}$ & AC increase $\geq 15 \text{ bpm}$
	Deceleration of FHR (DE): Onset from UC peak to FHR nadir $\geq 30 \text{ sec}$ EDE, No lag time ($DE < 1 \text{ sec}$) LDE, Lag time ($DE \geq 18 \text{ sec}$) Variable (DE): Onset from UC peak to FHR nadir $< 30 \text{ sec}$ and VDE decrease $\geq 15 \text{ bpm}$ (15-120 sec)
The Rules	<u>Category III</u>
	BLB and BVA $\geq 10 \text{ min}$
	BVA and LDE
	BVA and VDE
	<u>Category I</u>
	BLN and BVMO and EDE
	<u>Category II</u>
Exception of I & III	

407 **Table 2. Parameters of pregnant women and newborns**

	FD, N=30	Non-FD, N=30
Maternal age, years	33±9	32±7
Gestational age, weeks	30±9	32±8
Delivery		
Cesarean section	27	8
Non cesarean section	3	22
APGAR 1'		
<7	14	6
≥7	16	24
APGAR 5'		
<7	3	0
≥7	27	30
Anesthesia		
General anesthesia	2	3
Non general anesthesia	28	27
Fetus weight, g	2055±1371	2367±1771
Male fetus	12	20

408

409

410 **Table 3. Comparison results of clinicians and our rule-based method**

factor categories	Kappa (Selection number from clinicians' analysis and our rule-based method)
I	0.82 (36/30)
II	0.61 (12/23)
III	0.74 (12/7)
average	0.72

411 *there was no statistic difference between clinicians analysis and our rule-based method groups

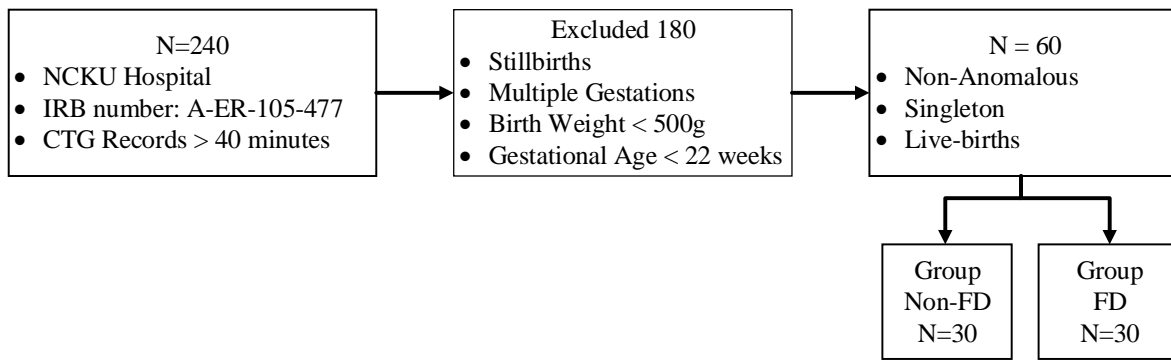
412

413

414 **Table 4. The results of FD ratio in each category between clinicians and the proposed**
 415 **method**
 416

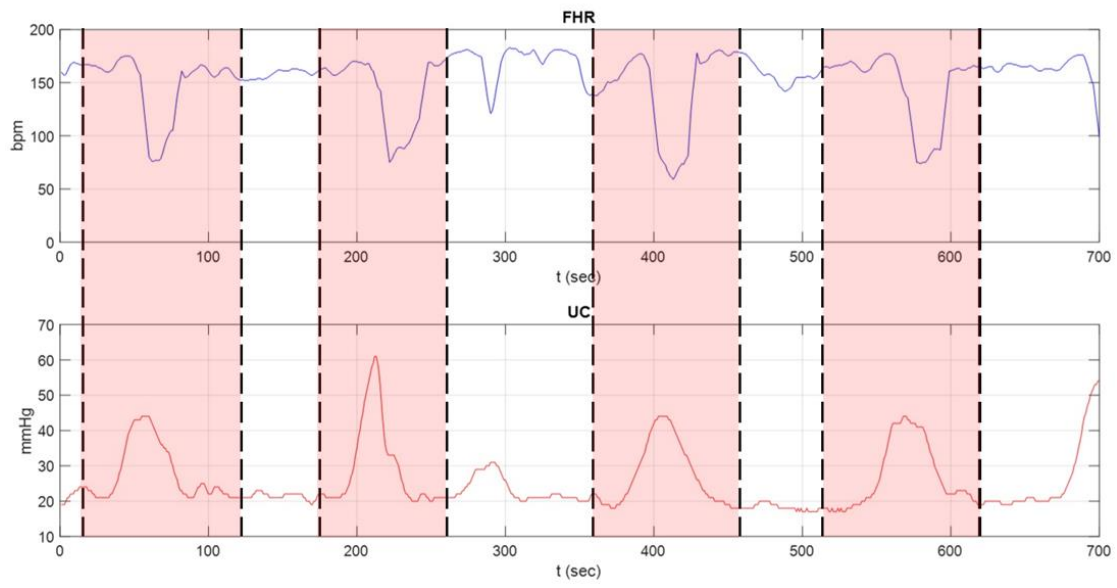
Categories \ Method		Clinicians' analysis	The proposed method
I		30.0%	22.8%
II	IIa	64.3%	28.8%
	IIb		71.2%
III		92.8%	88.8%

417
 418



419
420
421
422

Figure 1. The flowchart of clinical CTG selection.



423

424 **Figure 2. CTG segmentation based on UC analysis**