Received: 17 February 2023 | Revised: 20 August 2023 | Accepted: 3 September 2023

DOI: 10.1111/aogs.14680

AOGS Acta Obstetricia et Gynecologica Scandinavica

ORIGINAL RESEARCH ARTICLE

Is the fetus fit for labor? Introducing fast-and-frugal trees (FFTrees) to simplify triage of women for STAN monitoring: An interobserver agreement comparison with traditional classification

Susana Pereira[1] | Petra Bakker[2] | Ahmed Zaima[3] | Tullio Ghi[4] | Jörg Kessler[5] | Susanna Timonen[6] | Christoph Vayssière[7] | Katrin Löser[8] | Kaisa Holmberg[6] | Yves Jacquemyn[9] | Edwin Chandraharan[10] | David Wertheim[11] | Per Olofsson[12]

[1]Fetal Medicine Unit, The Royal London Hospital, Barts Health NHS Trust, London, UK

[2]Department of Obstetrics and Gynecology, Amsterdam UMC, Amsterdam, The Netherlands

[3]Department of Obstetrics and Gynaecology, Kingston Hospital NHS Foundation Trust, London, UK

[4]Department of Medicine and Surgery, University of Parma, Parma, Italy

[5]Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, and Department of Clinical Science, University of Bergen, Bergen, Norway

[6]Department of Obstetrics and Gynecology, Turku University Hospital, Turku, Finland

[7]Department of Obstetrics and Gynecology, Paule de Viguier Hospital, Toulouse III University, Toulouse, France

[8]Department of Obstetrics and Gynecology, South Jutland Hospital Aabenraa Campus, Aabenraa, Denmark

[9]Department of Obstetrics and Gynecology, Antwerp University Hospital, Edegem, and Global Health Institute, Antwerp University, Antwerp, Belgium

[10]Global Academy of Medical Education & Training, London, UK

[11]School of Computing and Information Systems, Faculty of Science, Engineering and Computing, Kingston University, London, UK

Abstract

Introduction: It is a shortcoming of traditional cardiotocography (CTG) classification table formats that CTG traces are frequently classified differently by different users, resulting in poor interobserver agreements. A fast-and-frugal tree (FFTree) flow chart may help provide better concordance because it is straightforward and has clearly structured binary questions with understandable "yes" or "no" responses. The initial triage to determine whether a fetus is suitable for labor when utilizing fetal ECG ST analysis (STAN) is very important, since a fetus with restricted capacity to respond to hypoxic stress may not generate STAN events and therefore may become falsely negative. This study aimed to compare physiology-focused FFTree CTG interpretation with FIGO classification for assessing the suitability for STAN monitoring.

Material and methods: A retrospective study of 36 CTG traces with a high proportion of adverse outcomes (17/36) selected from a European multicenter study database. Eight experienced European obstetricians evaluated the initial 40 minutes of the CTG recordings and judged whether STAN was a suitable fetal surveillance method and whether intervention was indicated. The experts rated the CTGs using the FFTree and FIGO classifications at least 6 weeks apart. Interobserver agreements were calculated using proportions of agreement and Fleiss' kappa (κ).

Results: The proportions of agreement for "not suitable for STAN" were for FIGO 47% (95% confidence interval [CI] 42%–52%) and for FFTree 60% (95% CI 56–64), ie a significant difference; the corresponding figures for "yes, suitable" were 74% (95% CI 71–77) and 70% (95% CI 67–74). For "intervention needed" the figures were 52% (95% CI 47–56) vs 58% (95% CI 54–62) and for "expectant management" 74% (95% CI 71–77) vs 72% (95% CI 69–75). Fleiss' κ agreement on "suitability for STAN" was 0.50 (95%

Abbreviations: CTG, cardiotocography; FFTree, fast-and-frugal tree; FHR, fetal heart rate; FIGO, International Federation of Gynecology & Obstetrics.; STAN, fetal ECG ST analysis.

[12]Institution of Clinical Sciences Malmö, Lund University, Malmö, Sweden

Correspondence
Susana Pereira, Fetal Medicine Unit, The Royal London Hospital, Whitechapel Road, London E1 1FR, UK.
Email: susana.pereira@nhs.net

CI 0.44–0.56) for the FIGO classification and 0.57 (95% CI 0.51–0.63) for the FFTree classification; the corresponding figures for "intervention or expectancy" were 0.53 (95% CI 0.47–0.59) and 0.57 (95% CI 0.51–0.63).

Conclusions: The proportion of agreement among expert obstetricians using the FFTree physiological approach was significantly higher compared with the traditional FIGO classification system in rejecting cases not suitable for STAN monitoring. That might be of importance to avoid false negative STAN recordings. Other agreement figures were similar. It remains to be shown whether the FFTree simplicity will benefit less experienced users and how it will work in real-world clinical scenarios.

KEYWORDS
cardiotocography, classification, clinical guidelines, fast-and-frugal tree, fetal monitoring, interobserver agreement, labor, midwifery, obstetrics, STAN

1 | INTRODUCTION

Efforts to improve the detection of intrapartum hypoxia resulted in development of the fetal ECG ST segment analysis (STAN). The STAN system continuously evaluates the capacity of the fetus to respond to hypoxic stress by identifying changes of the ECG ST interval, as displayed by an increase of the T/QRS ratio or biphasic ST interval shape. To assess such changes, an initial baseline T/QRS level must be determined while the fetal heart rate (FHR) shows no signs of preexisting hypoxia.[1] The initial triage to judge a fetus fit for labor is thus critical for the STAN system to be reliable. It has been shown in retrospective cardiotocography (CTG) assessments that about 30% of neonates with hypoxic ischemic encephalopathy show abnormal CTG patterns already at the start of recording, which would disqualify them from STAN monitoring.[2,3]

The CTG interpretation is the weakest link in the STAN system and recent efforts to improve and simplify the CTG classification have not solved the problem, and possibly made it worse: after an update of the Swedish national guidelines in 2017, inspired by the new International Federation of Gynecology & Obstetrics (FIGO) guidelines from 2015 (FIGO2015), the national incidences of birth acidemia and low Apgar scores have worsened.[4,5–7]

The traditional CTG classification systems are based on visual "pattern recognition" and clustering of the different FHR element categories (baseline, variability, accelerations, decelerations) into a fixed matrix.[4,8,9] However, a traditional fixed matrix may not fit all situations, and an alternative fetal physiology-oriented classification has been proposed.[10] The essence of a physiology-oriented approach is a genuine understanding of fetal pathophysiology and an awareness of the individual ability of every fetus to withstand the strain of labor. It focuses on interpretation of the FHR pattern over time in relation to the fetal condition instead of recognition of a certain pattern in a limited time window. Rosén et al. explored cases of false-negative STAN recordings to indicate metabolic acidosis and claimed that 12 of 18 cases could have been identified with a physiology-oriented approach.[11]

Key message

The interobserver agreement among expert obstetricians using fast-and-frugal trees with a physiological approach for CTG interpretation is higher than with the traditional FIGO classification system in rejecting cases for STAN monitoring.

Another weakness in CTG interpretation is the table format, where combinations of deviant patterns may result in different classifications by different users. The fast-and-frugal tree (FFTree) (Figure 1) is a classification chart for simple and speedy decision-making that is easy to understand, memorize, teach and execute.[12] Operating with little information in a straightforward flowchart structure with one question at a time and an answer needed before proceeding, results in a heuristic approach leading to mental shortcuts that ease the cognitive strain to take decisions.[13] The FFTree has recently been proposed for use in a physiology-oriented CTG interpretation approach, with simple and clearly structured binary questions with either "yes" or "no" answers (oral communication Birth Congress 2022). There are no direct questions regarding decelerations, but an emphasis on baseline and variability. The heuristic order of questions has been agreed on among the experts, with baseline being the first cue, because it has been reported to be the one with higher reproducibility and therefore the one likely to be easier and quicker to answer.

In the present study we applied the FFTree approach in a physiology-oriented CTG interpretation system to triage women suitable for STAN monitoring, and compared its interobserver agreement with that of traditional modified FIGO CTG interpretation matrix.[1] Due to its simplicity with a small number of cues, easiness to instruct, easy to remember and quick to execute,[13] we hypothesized that the FFTree chart would obtain better agreements than the FIGO classification when deciding whether a fetus is fit for

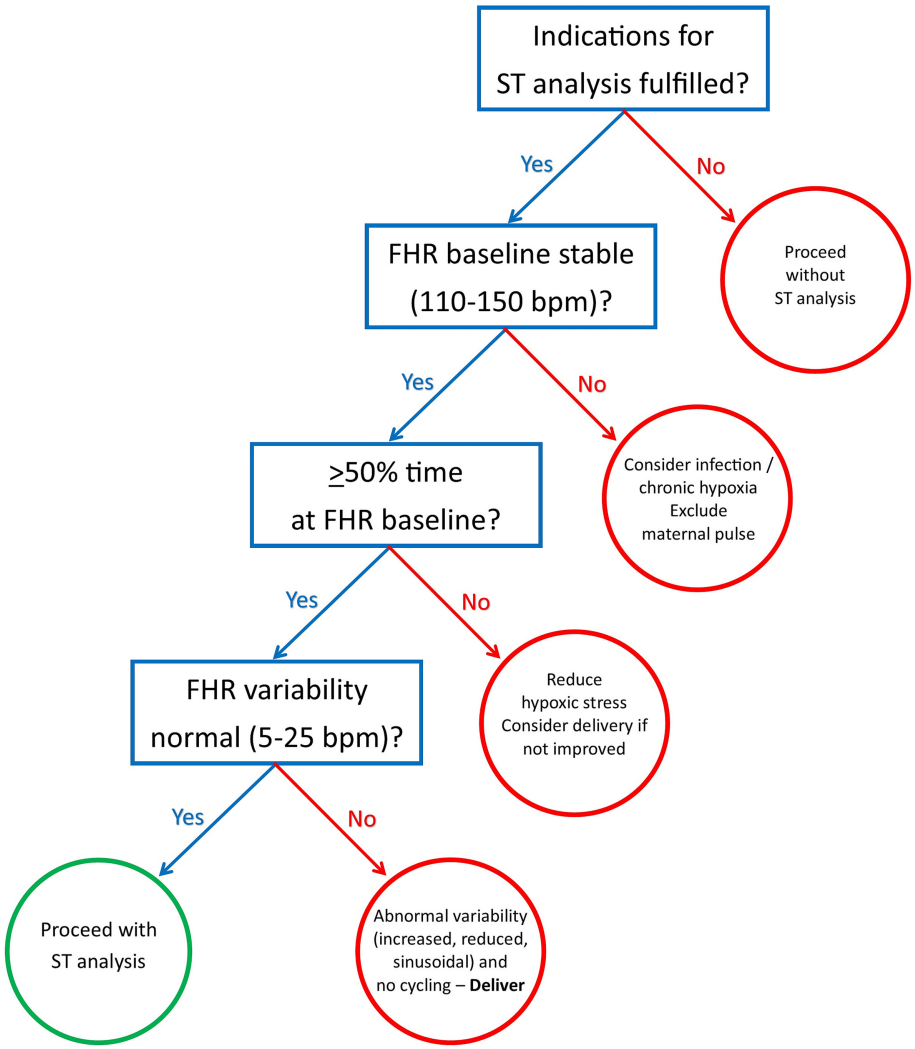

FIGURE 1 Fast-and-frugal tree to assess suitability for STAN monitoring at the beginning of CTG recording. Indications for ST analysis fulfilled refers to gestational age ≥36 weeks, no contraindications for fetal scalp electrode, not in active second stage of labor and informed consent obtained.

STAN monitoring, and whether the further management should be intervention or expectancy.

2 | MATERIAL AND METHODS

CTG traces from 36 women connected to STAN monitors in the first stage of labor were selected from a European multicenter study database of 6999 deliveries.[14] We intentionally selected a series with 50% adverse outcomes (low cord pH, low Apgar scores or intrapartum death) and the remaining with good outcomes but abnormal CTG traces according to FIGO classification (Table 1).[1]

Among the 36 cases, 17 had an adverse perinatal outcome (umbilical cord artery pH <7.0 and/or 5-minute Apgar score <7) with six instrumental deliveries, five cesarean sections and six spontaneous vaginal births. Among them, two intrapartum stillbirths occurred. In 19 cases the neonates had a cord pH ≥7.0 and 5-minute Apgar ≥7 after three instrumental deliveries, five cesarean sections and 11 spontaneous vaginal births.

Eight European experts on STAN monitoring (co-authors PB, TG, KH, YJ, JK, KL, ST, CV) were asked to evaluate the initial 40–60 minutes of the CTG recordings and judge whether STAN was a suitable fetal surveillance method and whether intervention (eg stop oxytocin augmentation, stop pushing, change position, deliver) was indicated or not. The judgments were based on the FFTree classification (Figure 1) in the first round and on the CTG classification used in the STAN algorithm in the second round (Table 1).[1] The CTG classification in the STAN algorithm is a modification of the FIGO classification from 1987 and the one currently recommended by the manufacturer of STAN monitors (Neoventa Medical AB, Mölndal, Sweden).[8] The round of CTG assessments with FIGO1987 was performed no earlier than 6 weeks after finishing the FFTree assessments. The experts were all specialists and consultants in obstetrics and gynecology with a median

TABLE 1 Cardiotocography (CTG) classification system used in the fetal ECG ST analysis (STAN) algorithm.[1]

CTG class	Baseline heart rate	Variability/reactivity	Decelerations
Normal	110–150 bpm	• 5–25 bpm • ≥2 accelerations/60 minutes	• No decelerations • Early uniform • Uncomplicated variable duration <60 seconds and beat loss <60 bpm
Suspicious	• 100–110 bpm • 150–170 bpm • Bradycardia <100 bpm for ≤3 min	• >25 bpm, salutatory • <5 for >40 minutes + absent accelerations	Uncomplicated variable duration <60 seconds and beat loss >60 bpm
	A combination of several suspicious patterns is classified as a pathological CTG		
Pathological	• 150–170 bpm + reduced variability • >170 bpm • Bradycardia <100 bpm for >3 min	• <5 bpm for >60 minutes • Sinusoidal pattern	• Complicated variable duration >60 seconds • Repeated late uniform • Combined
Preterminal	Absent variability (<2 bpm) + no accelerations, irrespective of baseline and decelerations		

of 20 years of obstetric practice (range 14–30 years), 14 years of which (range 5–21) were with STAN.

The experts (raters) were blinded to perinatal outcomes but not to gestational age, mode of onset of labor, oxytocin augmentation and ST analysis. The CTG traces were presented as screenshots from the Stan Viewer (Neoventa Medical AB) at a trace speed of 1 cm/min. A software program was specially developed for the study, where binary choice questions were electronically answered by the raters. There was no time limit to complete a case, but once a case was closed, no corrections could be made.

In contrast to the varying FIGO classifications, in the FFTree system there is no need of time thresholds beyond which the FHR patterns are classified as suspicious or pathological.[1,4,8] Periods of increased variability (bandwidth >25 bpm) <30 minutes were in the present study called zigzag pattern and periods lasting >30 minutes traditionally called saltatory pattern.[15,16]

Our FFTree system for STAN evaluation comprises three consecutive flow charts, where the first panel is for the initial triage assessment (Figure S1). It contains stepwise questions where a "no" answer should lead to action with escalation or multidisciplinary discussion being in itself a possible action. In the present study, only the initial triage assessments were studied (Figure 1).

2.1 | Statistical analyses

The raters' assessments of the 36 traces were stored on EXCEL (Microsoft Corporation) spreadsheet files. The data were analyzed using MINITAB v19 software (Minitab LLC) and software we developed in MATLAB v2022a (The MathWorks Inc.). MINITAB v19 software was used for assessing descriptive statistics and agreement with Fleiss' kappa. Graphs were prepared with Minitab and MATLAB. In addition, we prepared software in MATLAB to calculate proportions of agreement.[17] Agreement between raters was assessed using three separate methods: individual value point graphs, Fleiss' kappa (κ), and proportions of agreement.

Fleiss' kappa statistic with 95% confidence interval (95% CI) was used to assess the reliability of agreement between the eight raters. In comparisons between the classification systems, 95% CIs not overlapping each other indicated a significant difference. The kappa coefficient (κ) can adopt values between −1 and +1, where 0 represents the agreement that would be expected by chance and +1 reflects perfect agreement. The criteria of Landis & Koch were used to interpret the level of agreement from the κ value: $\kappa < 0$ poor, 0–0.2 slight, 0.21–0.4 fair, 0.41–0.6 moderate, 0.61–0.8 substantial, 0.81–1.0 almost perfect.[18]

The McNemar test was used to compare two correlated proportions, with a two-tailed $P < 0.05$ considered statistically significant.

2.2 | Ethics statement

The cases were selected from an anonymized European multicenter study database. Oral informed consent was obtained from all enrolled women at the time of data collection in the years 2000–2002, with due ethical approvals obtained in those maternity units where STAN was not standard care. Data from the raters were pseudo-anonymized for data analysis.

3 | RESULTS

3.1 | Suitability for commencing STAN monitoring

There was complete agreement between raters for suitability to commence STAN in 18 of 36 CTG traces (50%) using the FIGO classification and in 17/36 (47%) of traces using the FFTree classification (Table 2, Figure 2). Figure 2 shows the distribution of ratings "yes" and "no" to commence with STAN. For the FIGO classification the proportion of agreement for "no" was 47% (95% CI 42–52) and for "yes" 74% (95% CI 71–77). For the physiological FFTree approach the proportion of agreement for "no" was 60% (95% CI 56%–64%) and for "yes" 70%

TABLE 2 Agreements between eight expert raters on suitability for STAN and management of intrapartum fetal surveillance with fetal ECG ST segment analysis (STAN) ($n = 36$) as assessed by the traditional FIGO cardiotocography classification or the physiological FFTree classification. Figures are number of cases.

	Commence with STAN (yes/no)		Intervention or expectancy	
Management	FIGO	FFTree	FIGO	FFTree
Total agreement	18	17	15	18
One rater disagreeing	7	9	11	7
Two raters disagreeing	2	2	7	4
Three raters disagreeing	7	7	3	6
Four raters disagreeing	2	1	0	1

Abbreviations: FFTree, fast-and-frugal tree; FIGO, International Federation of Gynecology & Obstetrics.

(95% CI 67–74). The 95%CIs indicate a significantly higher agreement for the FFTree compared with FIGO for "no" but not for "yes".

When applying Fleiss' kappa statistics to the distribution of ratings in Figure 2 (eight raters, two categories), the κ agreement on suitability for STAN was 0.50 (95% CI 0.44–0.56) for the FIGO classification and 0.57 (95% CI 0.51–0.63) for the FFTree classification. The overlap of 95% CIs indicates there was no significant difference between the classifications, and both κ agreements were classified moderate.

3.2 | Agreements on intervention or expectant management

There was complete agreement on intervention or expectant management between the raters in 15/36 (42%) for the FIGO classification and in 18/36 (50%) for the FFTree classification (Table 2, Figure 3).

Figure 3 shows the distribution of "intervention" and "expectant management" ratings. For the FIGO classification the proportion of agreement for "intervention" was 52% (95% CI 47–56) and for "expectant management" 74% (95% CI 71–77). For the FFTree approach the proportion of agreement for "intervention" was 58% (95% CI 54–62) and for "expectant management" 72% (95% CI 69–75). The overlapping 95% CIs indicate no significant differences between the classifications.

Fleiss' kappa statistics applied on the distribution of ratings showed a κ agreement of 0.53 (95% CI 0.47–0.59) for the FIGO classification and 0.57 (95% CI 0.51–0.63) for the FFTree classification. The overlapping of 95%CIs indicates no significant difference between the classifications. Both κ values indicate moderate agreement.

3.3 | Details of maximum disagreement cases

Among the eight raters, total disagreements (4 vs 4 raters) were found in three cases (Figure 4). In cases nos 18 and 20, using the FIGO classification for judging suitability for STAN, the patterns were characterized by a shift towards increased variability (>25 bpm) after 20 minutes, a zigzag pattern. In case no. 23, with a total disagreement for suitability when using the FFTree classification, the FHR pattern was characterized by repetitive decelerations, with about 50% of time spent on the baseline. The only case of total disagreement regarding intervention was found in case no. 20 when using the FFTree classification.

3.4 | Analyses of cases with adverse outcomes

Seventeen of the 36 neonates had an adverse perinatal outcome as defined above. In 6/17 (35%) at least four operators in each case decided on intervention when using the FIGO classification, compared with 8/17 (47%) when using the FFTree classification ($P = 0.5$).

4 | DISCUSSION

This study showed moderate and consistent levels of interobserver agreement with the traditional FIGO CTG classification and the physiology-oriented FFTree classification in answering the question of whether the fetus is eligible for STAN monitoring and whether an intervention is indicated.[1] The FFTree classification had a significantly higher proportional agreement in rejecting cases for suitability of STAN monitoring, but for all other comparisons with proportional agreement and Fleiss' κ there were no such differences.

In this study, we compared two different classification systems using the same raters assessing the same CTG traces. There are numerous studies on inter- and intraobserver agreement in CTG classifications, but comparison of different observer agreements is not straightforward when performed with different methods, populations and circumstances. The κ value is dependent on the number of raters, categories, weighting applied and prevalence of the investigated outcome in the studied series; κ will be higher when there are fewer raters and categories and less difficult categorizations (less abnormality). If the prevalence of a positive rating is very high or very low, the chance agreement is high and κ is reduced accordingly.[19] We intentionally selected a series with 50% adverse outcomes (low cord pH, low Apgar scores, intrapartum death) and challenging CTG traces. It is thus not meaningful to compare κ values obtained in our study with κ values obtained in other studies because to compare two or more classification systems they should be applied to the same group of cases.

The concept of interobserver agreement is fundamental in the use of technology in medicine.[17] Poor intra- and interobserver agreements are well-known difficulties with CTG classification systems, as was already observed in the 1970s.[20,21] The FIGO1987 classification as well as other classification systems have persistently been criticized for low observer agreement. Thus, the FIGO guidelines from 1987 were updated in 2015 (FIGO2015) to make the interpretation simpler and more objective.[4] However, FIGO2015 was introduced without any prior scientific evaluation to show it was superior to the

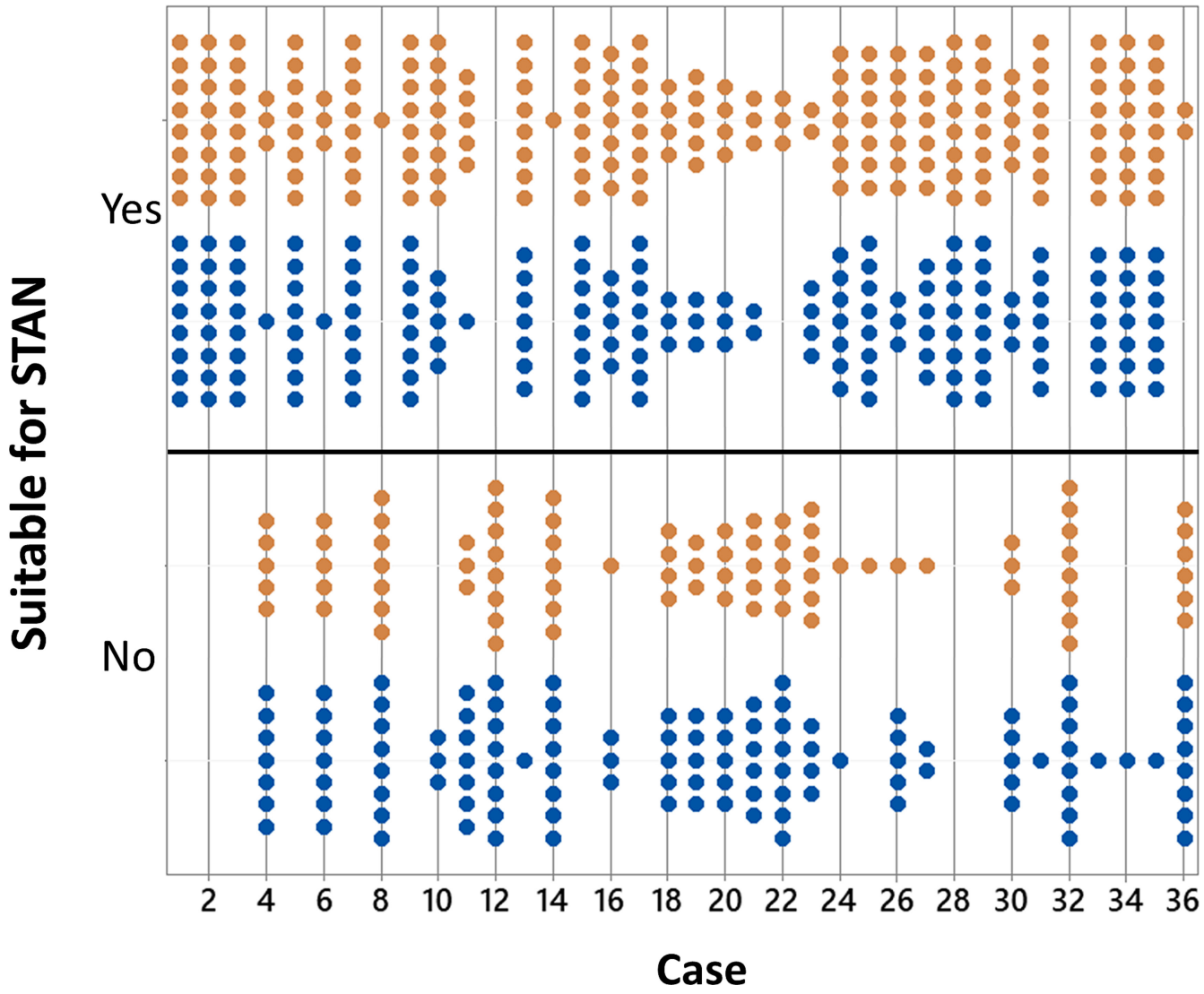

FIGURE 2 Suitability for commencing STAN monitoring. Distribution of "yes" (suitable) and "no" (not suitable) ratings when eight expert raters classified 36 CTG traces according to the FIGO classification (orange dots) and the fast-and-frugal tree classification (blue dots). Each dot represents one assessment in one case.

old but updated FIGO system from 1987.[1] Hitherto only two comparative outcome studies involving FIGO2015 have been published. The results of these studies are disappointing, showing a lower sensitivity for neonatal acidemia than its predecessor.[5,22] In the era of evidence-based medicine, comparative studies should be performed before rather than after the clinical implementation of a new classification, which was the principal reason for the present study.

In clinical practice the decision to commence STAN monitoring must be taken within a limited time frame, which in the present study was set to 40–60 minutes. In cases where the CTG patterns changed remarkably over a short period of time, the disagreement among raters was high; some raters might have overvalued the near normal pattern in the more recent time, and other raters might have focused on negative signs occurring at the end of the trace. Contrary to a saltatory pattern, a brief period of increased variability, called zigzag pattern, has traditionally not been considered a sign of fetal distress. However, recent studies demonstrate an association with low Apgar scores and low cord pH,[15,16] which may have misled some raters.

The eternal question of poor agreement between raters inspired Spilka et al. to make a more in depth analysis of the process of CTG evaluation.[23] They used the "latent class model" (LCM) instead of the simple majority (plurality) voting model, ie the aggregate evaluation of the raters. LCMs are typically used when a researcher suspects that subgroups or latent classes exist in a population. CTG classification templates have a fixed number of classes, but guidelines are not precise and are not strictly followed by clinicians, leaving room for alternative evaluations. The Spilka study showed that when using the original three-tier FIGO1987 classification system, the clinicians unconsciously used four classes. The difference between three and four classes was explained by a better separation of pathological traces, distinguishing a clearly pathological group with good agreement among clinicians.[23]

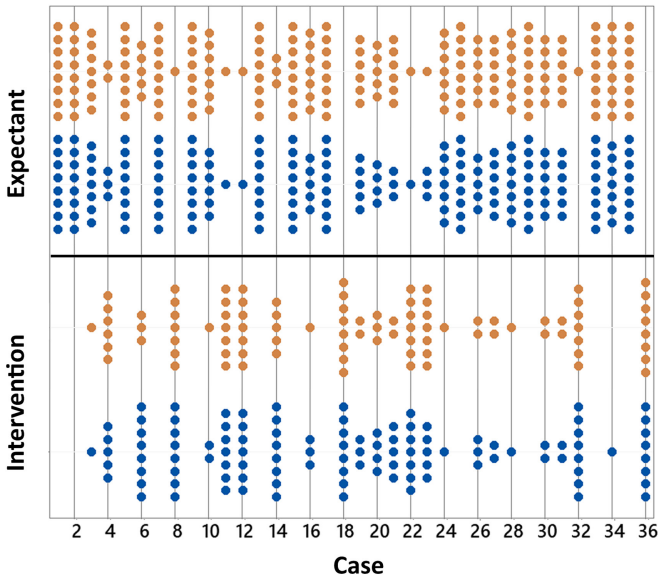

FIGURE 3 Distribution of "intervention" and "expectant management" ratings when eight expert raters classified 36 CTG traces according to the FIGO classification (orange dots) and the fast-and-frugal tree classification (blue dots). Each dot represents one assessment in one case.

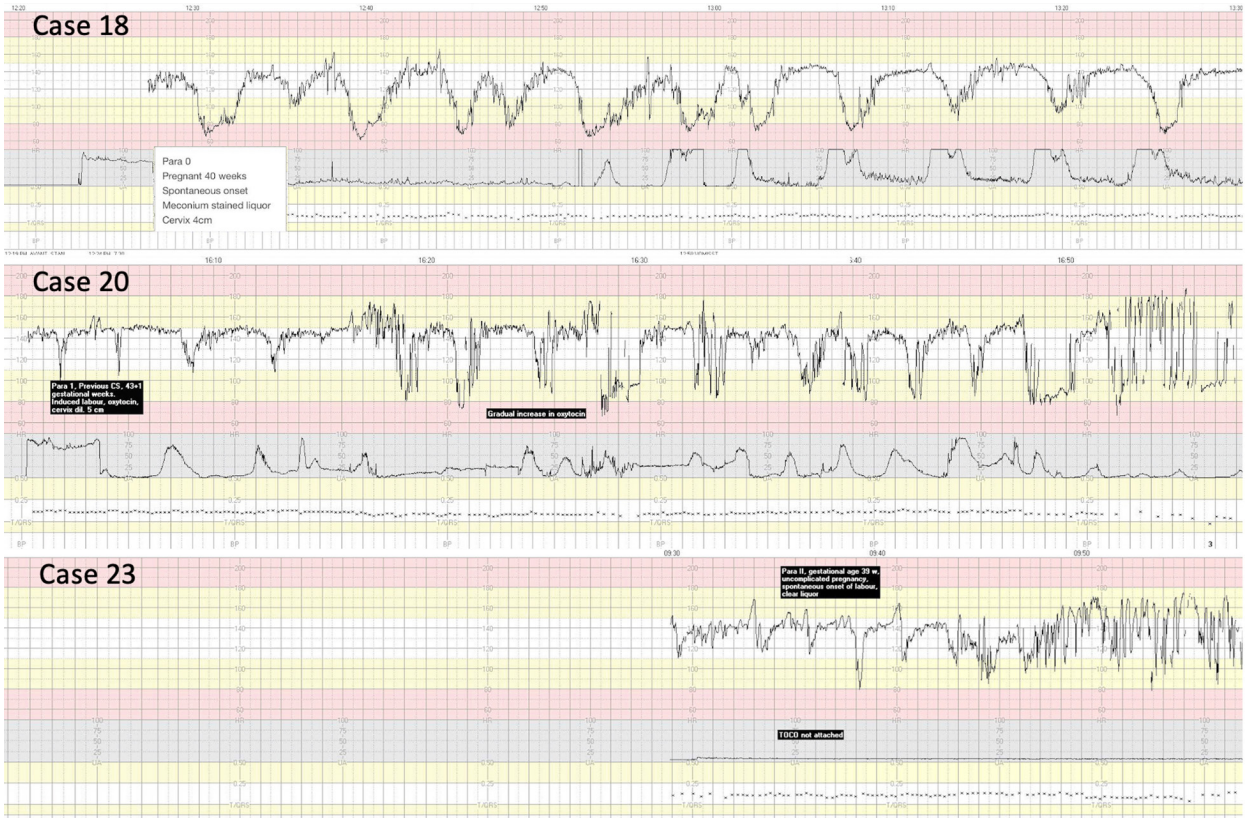

FIGURE 4 CTG traces with maximum disagreement between raters, referred to as nos 18, 20 and 23 in the text.

It is well described by behavioral scientists that when given a range of options from high to low, people tend to choose the middle option – this is known as the "extremeness aversion" phenomenon.[24] In the context of CTG interpretation, it might have an impact in situations of uncertainty, where the "suspicious" class becomes the popular middle choice. Since the FFTree classification is a heuristic system with all-inclusive and mutually exclusive binary options, the extremeness aversion phenomenon and a build-up of unconscious latent classes can be avoided.

The FFTree physiology-oriented classification fulfills the requirements of a simple structure with easily understandable definitions, enabling classifications to be made quickly. We observed that the

FFTree performs as well as the standard method when used by experts under no time pressure.

There are some limitations to our study. First, our results are valid for only the initial 40–60 minutes of monitoring. A CTG trace should also be evaluated in light of the fetal status and clinical situation, but in this type of retrospective study the clinical data that raters have access to, are always constrained. Finally, the unblinded T/QRS segment analysis might have influenced the management decision in cases where the CTG trace was particularly difficult to interpret. On the other hand, these limitations are expected to be equal in the two classification systems.

It was a strength of the study that we compared the classifications systems with two different statistical methods and displayed the results visually. The raters all had at least 14 years in obstetric practice experience, and with a minimum of 5 years using STAN. However, that might also be a limitation since in clinical practice, obstetricians and midwives with different levels of experience are involved in CTG interpretation. This should be considered in future studies. We also regard it as a strength that eight raters participated in the study. With fewer raters, the agreement figures would become less applicable in a clinical context.

5 | CONCLUSION

This is the first study on classification of CTG traces using an FFTree approach compared with a traditional CTG interpretation methodology. The FFTree physiology-oriented classification aims to provide a simple and easy to use structure.

The findings in the study show that when used with no time pressure, the proportion of agreement among experts using FFTrees physiological approach was significantly higher than with the traditional FIGO classification system in rejecting cases for STAN. This is important to avoid false-negative STAN recordings. In other aspects of agreement the systems were similar. It is yet to be explored whether its simplicity will be an advantage for less experienced users and in the real clinical scenario.

ACKNOWLEDGMENTS

We would like to express our sincere graitude to late Professor KG Rosen, who was a pioneer of STAN technology, for his advice and input into the designing of this study. We also thank Professor Gerd Gigerenzer for the fruitful discussions regarding heuristics at the beginning of this work.

AUTHOR CONTRIBUTIONS

SP: conceptualization, planning, designing study, collecting cases, recruiting raters, statistical analyses, analyzing results, literature search, writing the paper. PB: rater, analyzing results, literature search, writing the paper. AZ: creating the software questionnaire program, designing study. EC: conceptualization, designing study. DW: planning, designing study, statistical analyses, analyzing results, literature search, writing the paper. PO: designing study, analyzing results, literature search, writing the paper. TG, KH, YJ, JK, KL, ST, CV: raters, reviewing draft paper.

CONFLICT OF INTEREST STATEMENT

SP, AZ and YJ have received fees from Neoventa Medical AB, the manufacturer of STAN monitors, for education and training of obstetricians and midwives. PO is a former consulting Global Medical Adviser to Neoventa Medical. The remainin authors have stated explicitly that there are no conflicts of interest in connection with this article.

ORCID

Susana Pereira https://orcid.org/0000-0002-7167-9052
Edwin Chandraharan https://orcid.org/0000-0002-5711-8515
David Wertheim https://orcid.org/0000-0002-1672-3970
Per Olofsson https://orcid.org/0000-0002-0792-1393

REFERENCES

1. Amer-Wåhlin I, Arulkumaran S, Hagberg H, Marsál K, Visser G. Fetal electrocardiogram: ST waveform analysis in intrapartum surveillance. *BJOG*. 2007;114:1191-1193.
2. Murray DM, O' Riordan MN, Horgan R, Boylan G, Higgins JR, Ryan CA. Fetal heart rate patterns in neonatal hypoxic-ischemic encephalopathy: relationship with early cerebral activity and neurodevelopmental outcome. *Am J Perinatol*. 2009;26:605-612.
3. Pereira S, Patel R, Zaima A, et al. Physiological CTG categorization in types of hypoxia compared with MRI and neurodevelopmental outcome in infants with HIE. *J Matern Fetal Neonatal Med*. 2022;13:1-9.
4. Ayres-de-Campos D, Spong CY, Chandraharan E. FIGO intrapartum fetal monitoring expert consensus panel. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynecol Obstet*. 2015;131:13-24.
5. Ekengård F, Cardell M, Herbst A. Impaired validity of the new FIGO and Swedish CTG classification templates to identify fetal acidosis in the first stage of labor. *J Matern Fetal Neonatal Med*. 2022;35:4853-4860.
6. Olofsson P, Ekengård F, Herbst A. Time to reconsider: have the 2015 FIGO and 2017 Swedish intrapartum cardiotocogram classifications led us from Charybdis to Scylla? *Acta Obstet Gynecol Scand*. 2021;100:1549-1556.
7. Jonsson M, Söderling J, Ladfors L, et al. Implementation of a revised classification for intrapartum fetal heart rate monitoring and association to birth outcome: a national cohort study. *Acta Obstet Gynecol Scand*. 2022;101:183-192.
8. FIGO Subcommittee on Standards in Perinatal Medicine. Guidelines for the use of fetal monitoring. *Int J Gynecol Obstet*. 1987;25:159-167.
9. Intrapartum care for healthy women and babies. Clinical guideline [CG190]. Published: 03 December 2014. Last updated: 14 December 2022. nice.org.uk
10. Pinas A, Chandraharan E. Continuous cardiotocography during labour: analysis, classification and management. *Best Pract Res Clin Obstet Gynaecol*. 2016;30:33-47.
11. Rosén KG, Håkan Norén H, Carlsson A. FHR patterns that become significant in connection with ST waveform changes and metabolic acidosis at birth. *J Matern Fetal Neonatal Med*. 2019;32:3288-3293.
12. Wang Y, Luan S, Gigerenzer G. Modeling fast-and-frugal heuristics. *Psych J*. 2022;11:600-611.

13. Katsikopoulos KV, Şimşek Ö, Buckmann M, Gigerenzer G. *Classification in the Wild. The Science and Art of Transparent Decision Making*. The MIT Press; 2020.
14. Luttkus AK, Norén H, Stupin JH, et al. Fetal scalp pH and ST analysis of the fetal ECG as an adjunct to CTG. A multi-center, observational study. *J Perinat Med*. 2004;32:486-494.
15. Gracia-Perez-Bonfils A, Vigneswaran K, Cuadras D, Chandraharan E. Does the saltatory pattern on cardiotocograph (CTG) trace really exist? The ZigZag pattern as an alternative definition and its correlation with perinatal outcomes. *J Matern Fetal Neonatal Med*. 2021;34:3537-3545.
16. Tarvonen M, Hovi P, Sainio S, Vuorela P, Anderson S, Teramo K. Factors associated with intrapartum ZigZag pattern of fetal heart rate: a retrospective one-year cohort study of 5150 singleton childbirths. *Eur J Obstet Gynecol Reprod Biol*. 2021;258:118-125.
17. Grant JM. The fetal heart rate trace is normal, isn't it? Observer Agreement of Categorical Assessments. *Lancet*. 1991;26(337):215-218.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
19. Sim J, Wright CC. The kappa statistics in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257-268.
20. Ayres-de-Campos D, Bernardes J. Twenty-five years after the FIGO guidelines for the use of fetal monitoring: time for a simplified approach? *Int J Gynaecol Obstet*. 2010;110:1-6.
21. Trimbos JB, Keirse MJ. Observer variability in assessment of antepartum cardiotocograms. *BJOG*. 1978;85:900-906.
22. Ekengård F, Cardell M, Herbst A. Low sensitivity of the new FIGO classification system for electronic fetal monitoring to identify fetal acidosis in the second stage of labor. *Eur J Obstet Gynecol Reprod Biol X*. 2020;9:100120.
23. Spilka J, Chudáček V, Janků P, et al. Analysis of obstetricians' decision making on CTG recordings. *J Biomed Inform*. 2014;51:72-79.
24. Neumann N, Böckenholt U, Sinha A. A meta-analysis of extremeness aversion. *J Consumer Psychol*. 2016;26:193-212.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pereira S, Bakker P, Zaima A, et al. Is the fetus fit for labor? Introducing fast-and-frugal trees (FFTrees) to simplify triage of women for STAN monitoring: An interobserver agreement comparison with traditional classification. *Acta Obstet Gynecol Scand*. 2023;00:1-9. doi:10.1111/aogs.14680