

## Impaired validity of the new FIGO and Swedish CTG classification templates to identify fetal acidosis in the first stage of labor

Frida Ekengård , Monika Cardell & Andreas Herbst

To cite this article: Frida Ekengård , Monika Cardell & Andreas Herbst (2021): Impaired validity of the new FIGO and Swedish CTG classification templates to identify fetal acidosis in the first stage of labor, The Journal of Maternal-Fetal & Neonatal Medicine, DOI: [10.1080/14767058.2020.1869931](https://doi.org/10.1080/14767058.2020.1869931)

To link to this article: <https://doi.org/10.1080/14767058.2020.1869931>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 06 Jan 2021.



[Submit your article to this journal](#)



Article views: 719



[View related articles](#)



[View Crossmark data](#)



Citing articles: 3 [View citing articles](#)

## Impaired validity of the new FIGO and Swedish CTG classification templates to identify fetal acidosis in the first stage of labor

Frida Ekengård, Monika Cardell and Andreas Herbst

Department of Obstetrics, and Gynecology, Skåne University Hospital, Institution of Clinical Sciences Lund University, Lund, Sweden

### ABSTRACT

**Introduction:** Cardiotocography (CTG) is the main method of intrapartum fetal surveillance. In 2015 a new guideline was introduced by the International Federation of Gynecology and Obstetrics (FIGO), FIGO-15. In Sweden it was adjusted to SWE-17, replacing the previous national template, SWE-09. This study, conducted at one university hospital and one regional hospital in southern Sweden, evaluated the diagnostic validity of these three templates to detect fetal acidosis during the first stage of labor.

**Material and methods:** A total of 73 neonates with pH <7.1 in umbilical cord artery or vein at cesarean delivery during the first stage of labor were identified retrospectively. For each acidotic neonate, three non-acidemic neonates, with a pH  $\geq$ 7.2 in cord artery and vein, and Apgar scores  $\geq$ 9 at five and ten minutes, in all 219 neonates, were selected. The CTG tracings before birth in acidemic neonates, and tracings at the same cervical dilatation in the non-acidemic neonates, were independently assessed by three professionals from the obstetric staff, blinded to group and clinical data. Based on their categorizations of the included variables (baseline, variability, accelerations, decelerations and contraction rate), each CTG tracing was systematically classified according to the three templates. The sensitivity and specificity to identify acidemia by the classification pathological were determined for each template. Interobserver agreement in the assessments of tracings as pathological or not was analyzed, using free-marginal Kappa index.

**Results:** The sensitivity for patterns classified as pathological to identify acidemia was similar for FIGO-15 (71%) and SWE-17 (77%,  $p = .13$ ), and the specificity was 97% for both. SWE-09 had a significantly higher sensitivity (95%,  $p < .001$ ) albeit with a lower specificity (90%,  $p < .001$ ) than the other two templates. Among acidemic neonates, the fraction of tracings classified as normal was higher with SWE-17 (9.6%) than with SWE-09 (0%;  $p = .01$ ) and FIGO-15 (1.4%;  $p = .06$ ). For tracings from neonates with acidemia, agreement for three independent assessors was strong ( $\kappa$  0.85) with SWE-09, and weak for FIGO-15 ( $\kappa$  0.47), and SWE-17 ( $\kappa$  0.51). For tracings from neonates without acidemia, the agreement was almost perfect for FIGO-15 ( $\kappa$  0.91), strong with SWE-17 ( $\kappa$  0.90) and moderate with SWE-09 ( $\kappa$  0.78).

**Conclusions:** The ability of FIGO-15 and SWE-17 to identify fetal acidosis is considered insufficient. The combination of a high sensitivity and a high specificity makes SWE-09 the most discriminatory template during the first stage of labor.

### ARTICLE HISTORY

Received 21 August 2020  
Revised 18 December 2020  
Accepted 23 December 2020

### KEYWORDS

Electronic fetal monitoring; fetal heart rate; cardiotocography; classification; acidosis



### Introduction

Cardiotocography (CTG) is the gold standard for intrapartum fetal surveillance even though randomized trials show modest benefits [1–4]. The primary goal of reducing the burden of perinatal mortality and long-term sequelae while avoiding unnecessary obstetric intervention has proven difficult to achieve [4–8]. Still, large population based studies have indicated that the use of

electronic fetal monitoring (EFM) is associated with decreased early neonatal mortality and morbidity [9].

Many efforts have been made to improve the efficacy of EFM, resulting in national guidelines with some variations in interpretation of the different CTG variables.

In 2015 the International Federation of Gynecology and Obstetrics (FIGO), published a new guideline on how to use CTG with a new interpretation template,

**CONTACT** Frida Ekengård  [Frida.ekengard@med.lu.se](mailto:Frida.ekengard@med.lu.se)  Department of Obstetrics and Gynecology, Skåne University Hospital, Institution of Clinical Sciences Lund University, Malmö, S-20502, Sweden

 Supplemental data for this article is available online at <https://doi.org/10.1080/14767058.2020.1869931>.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

FIGO-15 [10]. It was the first update from FIGO since 1987 [11]. The main objective behind this revision was to increase the effectiveness of EFM by creating objective definitions of the different features of CTG and increase intra- and interobserver agreement. The new template is a three-tier system, created by an international consensus panel [10].

In 2015 a national template was in use in Sweden since 2009, SWE-09 [12]. It was a version of the FIGO template from 1987 [11] adjusted to accommodate the fetal ECG ST analysis (STAN) algorithm from 2007 [13]. A modification of the FIGO-15 was introduced in Sweden during 2017, SWE-17 [14], and is now the national standard interpretation template. The SWE-17 was introduced in a wish to adhere to the new international guideline, and with the consideration that a low specificity for the SWE-09 template might cause unnecessary interventions. In Denmark, it was decided to await evidence before adopting the new FIGO guideline [15].

All three templates include the parameters baseline fetal heart rate, variability and decelerations. The SWE-09 also include the parameters accelerations and contraction rate. The two new templates accept a wider range of normal baseline heart rate than SWE-09 and differ in their classification of decelerations in that they stipulate a minimum frequency and duration for most decelerations in order to be deemed pathological. The three templates are summarized in [Figure 1](#).

Since neither of these templates had been evaluated before being taken in clinical use, and since CTG patterns are often markedly different in the first and second stage of labor, we planned two studies evaluating the sensitivity and specificity in the first and second stage of labor, respectively. The present study focuses on the first stage.

During the first stage of low-risk labor, Swedish guidelines recommend intermittent CTG monitoring every second hour and auscultation every 15–30 min in between [14], which has been shown to be as safe as continuous fetal monitoring in low-risk labor [16]. If the tracing is not classified as normal, extended monitoring is recommended. Hence, from low-risk labors with normal fetal heart rates, relatively short CTG tracings are recorded in the first stage of labor. Continuous CTG is recommended in high-risk labor [14].

The primary objective of this study was to compare the templates, SWE-09, FIGO-15 and SWE-17, regarding sensitivity and specificity in identifying acidosis during the first stage of labor.

The null hypothesis was that sensitivity and specificity do not differ between the three classification systems.

## Materials and methods

This study includes intrapartum CTG tracings recorded at Helsingborg Hospital, between March 13th 2012 and December 31st 2016 and at Skåne University Hospital, between April 23<sup>d</sup>2013 and October 31st 2017, Region Skåne, Sweden.

Neonates born by emergency cesarean section during the first stage of labor and having an umbilical cord arterial or venous pH <7.10 were identified. The first stage of labor was defined as regular contractions and a cervical dilation of three to nine centimeters. Further inclusion criteria included singleton pregnancy, gestational age  $\geq 34 + 0$  weeks and an available CTG tracing >15 min prior to delivery, with a maximum delay of 30 min from the end of the CTG tracing and delivery.

Normally, fetal pH is higher during the first stage of labor than at birth, since fetal pH declines and lactate increases during the second stage of labor [17,18]. A cord artery pH <7.10 has been associated with an increased risk of adverse neurological outcomes [19] and was therefore the cutoff used in this study, as indicator of exposure to hypoxia.

For each neonate born with acidemia, the first three neonates born after the acidemic neonate at the same unit and fulfilling the inclusion criteria were included. Inclusion criteria for the non-acidemic group were singleton birth  $\geq 34 + 0$  weeks, pH  $\geq 7.20$  in cord artery and vein, and Apgar score  $\geq 9$  at five and ten minutes. Furthermore, the non-acidemic neonates had to have an available CTG tracing for at least 15 min at the same cervical dilation as its corresponding acidemic neonate. Thus, non-acidemic neonates were matched to the acidemic neonates for the evaluation of CTG at the same cervical dilation, i.e. only including tracings from the first stage.

A total of 57,582 neonates were born at the two hospitals during the study period. Of 1470 neonates, born with an umbilical cord arterial or venous pH <7.10, 126 (8,6%) were delivered by cesarean in the first stage of labor. The study included 73 acidemic neonates and 219 non-acidemic neonates fulfilling the inclusion criteria.

Background data are shown in [Table 1](#). The acidemic group included more high-risk pregnancies. There were three preterm neonates, born at 34 + 5, 36 + 1 and 36 + 5 weeks, all in the non-acidemic group.

CTG tracings between 15 and 80 min were assessed and saved as anonymous PDF-files. The short lower limit of 15 min was chosen because of the use of intermittent CTG during first stage of labor. The graph

	SWE-09 <sup>a</sup>	FIGO-15 <sup>b</sup>	SWE-17 <sup>c</sup>
<b>Normal</b>			
Baseline (bpm)	110-150	110-160	110-160
Accelerations	≥ 2 per hour		
Variability (bpm)	5 - 25	5 - 25	5 - 25
Decelerations	None <i>or</i> uniform early <i>or</i> variable < 30 seconds	No repetitive <sup>1</sup>	No repetitive <sup>1</sup> <i>or</i> uniform early <i>or</i> variable < 1 min
Contraction rate	5 or less per 10 min		
<b>Suspicious</b>			<sup>2</sup>
Baseline (bpm)	100-110 <i>or</i> 150-170 <i>or</i> < 100 ≤ 3 min	Lacking at least one characteristic of normality but with no pathological features	100-109 <i>or</i> > 160
Accelerations	< 2 per hour		
Variability (bpm)	> 25 <i>or</i> < 5 > 40 min without accelerations		
Decelerations	Variable 30-60 seconds <i>or</i> amplitude > 60 beats		Repetitive <sup>1</sup> variable > 1 min with normal baseline and variability
Contraction rate	> 5 per 10 min		
<b>Pathological</b>			<sup>3</sup>
Baseline (bpm)	< 100 > 3 min <i>or</i> > 170	< 100	< 100
Variability (bpm)	< 2 <sup>4</sup> <i>or</i> < 5 > 60 min without accelerations <i>or</i> sinusoidal	< 5 > 50 min <i>or</i> > 25 > 30 min <i>or</i> Sinusoidal > 30 min	< 2 <i>or</i> < 5 > 60 min <i>or</i> > 25 > 30 min <i>or</i> Sinusoidal > 30 min
Decelerations	Variable > 1 min <i>or</i> uniform late <i>or</i> combined	Repetitive <sup>1</sup> late <i>or</i> prolonged > 3 min for > 30 min <i>or</i> > 20 min if reduced variability <i>or</i> one prolonged > 5 min	Repetitive <sup>1</sup> late > 30 min <i>or</i> Repetitive <sup>1</sup> late > 20 min with tachycardia <i>or</i> reduced variability <i>or</i> Repetitive <sup>1</sup> variable > 1 min with tachycardia <i>or</i> reduced variability > 20 min <i>or</i> Repetitive <sup>1</sup> prolonged > 3 min <i>or</i> One prolonged > 5 min

<sup>a</sup> National template that was in use in Sweden 2009-2017 (12)

<sup>b</sup> New template introduced by FIGO in 2015 (10)

<sup>c</sup> National template used in Sweden since 2017, modification of FIGO-15 (14)

Abbreviations; min; minute, bpm; beats per minute,

<sup>1</sup> repetitive; associated with more than 50% of uterine contractions

<sup>2</sup> Swe-17; a pattern that does not fulfil the criteria for normal or pathological, should be classified as suspicious

<sup>3</sup> Swe-09; two or more criteria for suspicious may lead to classification pathological

<sup>4</sup> Swe-09; classified as "preterminal"

**Figure 1.** Criteria of the three interpretation templates for the different classifications.

scale used was one centimeter per minute. Clinical data was gathered in Excel®.

The median duration of the CTG tracings in the academic group was 70 min and six tracings (8%) were <30 min. In the non-academic group, the median duration was 40 min and 68 tracings (31%) were <30 min. The total range was 16 to 80 min in both groups.

In 20 (27%) of the academic neonates, scalp blood sampling had been performed within 30 min before the end of the CTG-tracing. Of those, 14 had lactate

values above the cutoff for intervention (4.8 mmol/L) for the device Lactate Pro® used during the study period. The cutoff represented the 75th percentile in a study by Kruger *et al.* [20].

The anonymized tracings were numbered and arranged by computer randomization. Each tracing was interpreted by three assessors (midwives and physicians) for whom the interpretation of CTG was part of their daily work at a labor ward. The only information available to the assessors was that the CTG

**Table 1.** Summary of background data of acidemic neonates, pH <7.1 in umbilical cord artery or vein, after delivery with cesarean section, and in non-acidemic neonates with pH  $\geq$ 7.2 in both cord vessels and Apgar scores  $\geq$ 9 at five and ten minutes.

	Acidemic neonates n (%)	Non-acidemic neonates n (%)
Total	73	219
Primiparous	34 (46.6)	87 (39.7)
Cesarean delivery <sup>a</sup>	73 (100)	20 (9.1)
Instrumental delivery	–	11 (5.0)
Oxytocin augmentation <sup>b</sup>	25 (34.2)	93 (42.5)
Preterm birth <37 + 0	0 (0)	3 (1.4)
Postterm birth $\geq$ 42 + 0	8 (11.0)	23 (10.5)
SGA <sup>c</sup>	6 (8.2)	8 (3.7)
LGA <sup>d</sup>	6 (8.2)	5 (2.3)
Breech	0 (0)	0 (0)
Fever/infection	6 (8.2)	8 (3.7)
Meconium stained amniotic fluid	32 (43.8)	48 (21.9)
Diabetes	8 (11.0)	9 (4.1)
Preeclampsia	6 (8.2)	5 (2.3)
Body mass index <25 kg/square meter	30 (41.0)	121 (55.2)
Body mass index >30 kg/square meter	19 (26.0)	21 (9.6)
Smoking	5 (6.8)	18 (8.2)
Female fetus	26 (35.6)	106 (48.4)
5-minute Apgar score <7 <sup>e</sup>	15 (20.5)	–
Base excess <sup>f</sup> <-12 mmol/L	17 (23.3)	0 (0)

<sup>a</sup>Inclusion criteria for acidemic neonates.

<sup>b</sup>At any time during labor.

<sup>c</sup>Small from gestational age-specific mean weight, defined as birthweight < -2 SD from expected.

<sup>d</sup>Large from gestational age-specific mean weight, defined as birthweight > +2 SD from expected.

<sup>e</sup>Apgar scores  $\geq$ 9 at five and ten minutes were inclusion criteria for non-acidemic neonates.

<sup>f</sup>Standard base excess, calculated for the extracellular compartment by the blood-gas analyzer (Radiometer<sup>TM</sup>), data incomplete in 21 acidemic neonates and 4 non-acidemic neonates.

**Table 2.** Classification of 292 cardiotocograms (CTG) according to three different classification templates (see footnote for explanation).

	SWE-09 <sup>a</sup> n (%)	FIGO-15 <sup>b</sup> n (%)	SWE-17 <sup>c</sup> n (%)
Acidemic neonates, N	73	73	73
CTG class			
Normal	0 (0)	1 (1.4)	7 (9.6)
Suspicious	4 (5.5)	20 (27.4)	10 (13.7)
Pathological	69 (94.5)	52 (71.2)	56 (76.7)
Non-acidemic neonates, N	219	219	219
CTG class			
Normal	166 (75.8)	165 (75.3)	204 (93.2)
Suspicious	31 (14.2)	47 (21.5)	8 (3.7)
Pathological	22 (10.0)	7 (3.2)	7 (3.2)

Classifications were first performed independently by three staff professionals and then the CTG class was determined by majority decision.

<sup>a</sup>National template that was in use in Sweden 2009–2017 [12].

<sup>b</sup>New template introduced by FIGO in 2015 [10].

<sup>c</sup>National template used in Sweden since 2017, modification of FIGO-15 [14].

tracings were from singleton pregnancies, from the first stage of labor and that all neonates were born with either a low or a normal cord blood pH.

The assessors completed forms with their interpretation of the different CTG parameters individually (Supplement 1 and 2).

The final classification according to each template, utilizing the interpretation of the included parameters made by the assessors, was then made by one of the authors (FE). The final classification was according to the majority. If the comprehensive CTG class differed

between all three interpreters, a senior obstetrician made a fourth interpretation of the variables to receive a majority interpretation. Thus, this interpretation worked as a casting vote.

The assessors were blinded to group and other clinical variables, as were the authors until the analyses were complete. The SWE-09 is a four-tier template. The fourth classification is preterminal, a loss of variability. In the study pathological and preterminal were merged to allow statistical comparisons with the two three-tier templates. For the final analysis, a CTG classified as pathological was considered a positive test, whereas a normal or suspicious CTG was considered as a negative test

### Statistical analyses

Stat View<sup>®</sup> computer software (SAS Institute, version 5.0.1; Cary, NC) was used to gather and analyze the data. Sensitivity and specificity with 95% confidence intervals (CI) were calculated, using [www.sample-size.net/confidence-interval-proportion](http://www.sample-size.net/confidence-interval-proportion) provided by the University of California, San Francisco. A two-sided McNemar's test was also used to determine the statistical significance of differences in classification of the CTG tracings with the different templates. A p-value <.05 was considered as statistically significant.

**Table 3.** Comparison of sensitivity and specificity for the different templates to identify neonatal acidosis in the first stage of labor.

	SWE-09 <sup>a</sup>	FIGO-15	SWE-17
Sensitivity, % (95% CI)	94.5 (86.6–98.5) <sup>b,c</sup>	71.2 (59.4–81.2) <sup>b,d</sup>	76.7 (65.4–85.8) <sup>c,d</sup>
Specificity, % (95% CI)	90.0 (85.2–93.6) <sup>e,f</sup>	96.8 (93.5–98.7) <sup>e,g</sup>	96.8 (93.5–98.7) <sup>f,g</sup>

<sup>a</sup>Preterm and pathological patterns merged to one class.

<sup>b</sup>McNemar's test; SWE-09 vs FIGO-15  $p = .0001$ .

<sup>c</sup>McNemar's test; SWE-09 vs SWE-17  $p = .0009$ .

<sup>d</sup>McNemar's test; FIGO-15 vs SWE-17  $p = .13$ .

<sup>e</sup>McNemar's test; SWE-09 vs FIGO-15  $p = .0007$ .

<sup>f</sup>McNemar's test; SWE-09 vs SWE-17  $p = .0003$ .

<sup>g</sup>McNemar's test; FIGO-15 vs SWE-17  $p = .48$ .

**Table 4.** Comparison of inter-observer agreement in classifications of CTG patterns between three independent interpreters using the same classification chart.

	SWE-09	FIGO-15	SWE-17
<b>Acidemic neonates, N</b>	73	73	73
Overall agreement	92.7%	73.5%	75.3%
<b>Free-marginal kappa, <math>\kappa</math> (95% CI)</b>	0.85 (0.80–0.95)	0.47 (0.32–0.62)	0.51 (0.36–0.66)
<b>Non-acidemic neonates, N</b>	219	219	219
Overall agreement	89.0%	95.7%	95.1%
<b>Free-marginal kappa, <math>\kappa</math> (95% CI)</b>	0.78 (0.72–0.85)	0.91 (0.87–0.96)	0.90 (0.86–0.95)

For each chart the classification could be normal, suspicious or pathological. "Preterminal" in SWE-09 was included in the classification pathological. Acidemic neonates; pH <7.10 in cord artery or vein, after delivery with cesarean section in the first stage of labor. Non-acidemic neonates; pH  $\geq 7.20$  in both cord vessels and Apgar scores  $\geq 9$  at five and ten minutes. Free-marginal kappa ( $\kappa$ ) according to Justus Randolph (with 95% confidence intervals) were calculated for assessment of tracings as pathological versus not pathological (normal and suspicious).

For agreement, a free-marginal kappa index for multiple raters was calculated according to J. Randolph, using an online Calculator, [www.justusrandolph.net/kappa](http://www.justusrandolph.net/kappa) and McHugh [21] was used for classification of kappa index.

### Ethical approval

The Regional Ethical Review Board in Lund gave ethical approval, Dnr 2016/371, 2016-05-24.

### Results

The final classifications are summarized in Table 2, and the sensitivity and specificity for the three templates in Table 3. The classification pathological in SWE-09 had a significantly higher sensitivity for identifying acidotic neonates (94.5%) than both SWE-17 (76.7%;  $p = .0009$ ), and FIGO-15 (71.2%;  $p = .0001$ ). The specificity was significantly lower for SWE-09 (90.0%) than for the other two templates (96.9% for both; FIGO-15;  $p = .0007$  and SWE-17;  $p = .0003$ ).

For tracings from acidemic neonates, agreement between the three assessors was strong for SWE-09 ( $\kappa$  0.85), and weak with the new templates, FIGO-15 ( $\kappa$  0.47), and SWE-17 ( $\kappa$  0.51). For the non-acidemic group the agreement was almost perfect for FIGO-15 ( $\kappa$  0.91), strong with SWE-17 ( $\kappa$  0.90), and moderate for SWE-09 ( $\kappa$  0.78), Table 4.

The most common patterns among acidemic neonates that were classified as pathological with SWE-09 but not with the two other templates were combinations of: complicated variable decelerations ( $n = 13$ ), combined decelerations ( $n = 9$ ), uniform late decelerations ( $n = 5$ ), tachycardia ( $n = 9$ ), decreased variability ( $n = 6$ ) and lack of accelerations ( $n = 10$ ).

In the acidemic group, the fraction of tracings classified as normal was higher with SWE-17 (9.6%) than with SWE-09 (0%;  $p = .01$ ) and FIGO-15 (1.4%;  $p = .06$ ).

In the group of non-acidemic neonates a higher fraction of the traces was classified as pathological with SWE-09 (10.0%) than with the SWE-17 and FIGO-15 (3.2%;  $p < .01$  for both). The most common pattern only classified as pathological with SWE-09 was complicated variable decelerations ( $n = 11$ ).

### Discussion

For safe intrapartum care, an interpretation template for fetal monitoring should identify a high rate of neonates with acidosis, especially during the first stage of labor. In our study, SWE-09 demonstrated the highest sensitivity (95%) for the classification pathological. FIGO-15 and SWE-17 showed lower sensitivity to detect acidotic neonates (71%,  $p = .0001$  and 77%,  $p = .0009$  respectively) and therefore, do not provide safe guidance if classification pathological is the only class used to indicate fetuses at risk for whom intervention should be considered.

The higher rate of patterns classified as normal in acidotic cases (9.6%) is also a matter of concern for SWE-17, since a normal classification can lead to that the CTG-recording is terminated, leaving a fetus already exposed to oxygen deficit without continuous surveillance.

When CTG is used as a screening method, with an option of fetal scalp blood sampling as a secondary test if the pattern is pathological, the specificity for the template is not as important as the sensitivity, since the secondary test will increase specificity. Moreover, since the purpose of intrapartum fetal monitoring is to prevent rather than to predict fetal acidosis, a specificity for fetal acidosis close to 100% may not be achievable. We consider that the specificity of 90% for pathological patterns with SWE-09 may indicate a clinically useful level.

The new templates had higher interobserver agreement in tracings from non-acidemic neonates, but lower interobserver agreement for acidemic neonates. This is in line with the results for sensitivity and specificity, and we speculate that SWE-09 provides better tools to identify abnormality, whereas FIGO-15 and SWE-17 may include better tools to identify normality.

The probability in identifying an abnormal CTG pattern is probably lower during a short than during a long CTG tracing. The shorter median duration of CTG tracings for non-acidemic neonates than for acidemic neonates might have led to higher specificity than if the tracings had been of identical duration. This, however, would affect all templates and not the comparison between the templates.

In the group of acidemic neonates a time gap of up to 30 min between the end of the CTG tracing and delivery was allowed. It is possible that some acidemic neonates might have become acidotic during this gap, which would have decreased sensitivity for all templates.

Three previous studies, comparing FIGO-15 to older templates are published. Olofsson et al. compared the FIGO-15, the SWE-17 and the STAN interpretation algorithm, similar to SWE-09, and found discrepancies in the classification of CTG with the three templates [22] and that FIGO-15 has a lower sensitivity than the STAN interpretation algorithm [23]. Marti Gamboa et al also found a low sensitivity for FIGO-15, similar to the five-tier system by Parer and Ikeda [24]. Our study confirms the low sensitivity for FIGO-15, and is to our knowledge the first to compare the new FIGO template with other templates specifically during the first stage of labor.

The results of the present study and the three previous, support holding on to previous templates based

on the FIGO guidelines from 1987 [11], until solid evidence for new guidelines have been presented.

A national change in Sweden during 2017 to a template with a lower sensitivity might be reflected in national data. The annual report from the Swedish Pregnancy Register, reported an increase in the rate of 5-min Apgar scores <7, from 1.0% in 2016 to 1.2% in 2019 [25]. The Swedish Pregnancy Register covers over 90% of Swedish deliveries, and has been validated [26].

Due to the worrying increase in low Apgar scores, we retrieved data from the Swedish Pregnancy Register regarding 5-min Apgar scores in term live births, elective cesarean deliveries excluded, for 2014–2016 ( $n = 264,181$ ), and 2018–2019 ( $n = 180,120$ ). In these cohorts, the rate of 5-min Apgar scores <4 increased from 0.18% during 2014–2016 to 0.25% during 2018–2019 ( $p < .0001$ ). Whether this has to do with the change of classification template or with other factors is unknown and must be further studied. The rate of emergency cesarean sections among live term births, elective cesareans excluded, was similar during 2014–2016 (9.1%), and 2018–2019 (8.9%;  $p = .06$ ), whereas instrumental deliveries decreased from 6.4% to 5.5% ( $p < .0001$ ).

A recent study showed a potential benefit of including clinical risk factors into a CTG interpretation template [27]. In our study the midwives and physicians classifying CTG were blinded to outcome to eliminate the risk of ascertainment bias [28,29]. Many clinical factors other than the CTG pattern are important for decision making, as well as a physiological understanding of the fetal heart rate changes [30,31], but this was not the focus of our study.

Of the three studied templates, the SWE-09 is the most demanding for the classification normal, with an upper normal baseline heart rate limit of 150 bpm, a requirement of accelerations, and it limits the acceptance of late or complicated variable decelerations. This may explain the superior sensitivity for acidemia. SWE-17 and FIGO-15 have higher demands for the classification pathological, requiring a minimum duration of repetitive late and complicated variable decelerations, which may explain higher specificity. Further studies of the separate variables included in the templates are planned. We will continue our research with a goal to optimize a template to achieve the highest possible validity in diagnosing fetal hypoxia with CTG.

## Conclusion

For FIGO-15 and SWE-17 the sensitivity for pathological patterns to identify neonates with fetal acidosis

is considered insufficient. SWE-09 is considered the most discriminatory template during the first stage of labor, combining a high sensitivity and a high specificity for pathological patterns. With the current knowledge, we would recommend adhering to previous guidelines emanating from FIGO guidelines from 1987, until any adjustments have been scientifically evaluated.

## Acknowledgements

We gratefully acknowledge all midwives and physicians who classified all these CTG tracings.

## Disclosure statement

The authors declare no competing financial interest that have influenced the work reported in this paper. Andreas Herbst has contributed in the development of the Swedish classification systems from 2009 and 2017.

## Funding

This work was supported by research grants from Region Skåne and funding from LÖF, the national Swedish patient insurance company. The funders played no role in planning or conducting the research or writing of the paper.

## References

- [1] MacDonald D, Grant A, Sheridan-Pereira M, et al. The Dublin randomized controlled trial of intrapartum fetal heart rate monitoring. *Am J Obstet Gynecol.* 1985;152(5):524–539.
- [2] Vintzileos AM, Antsaklis A, Varvarigos I, et al. A randomized trial of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation. *Obstet Gynecol.* 1993;81(6):899–907.
- [3] Vintzileos AM, Nochimson DJ, Antsaklis A, et al. Comparison of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation in detecting fetal acidemia at birth. *Am J Obstet Gynecol.* 1995;173(4):1021–1024.
- [4] Alfirevic Z, Devane D, Gyte GM, et al. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev.* 2017;2:CD006066.
- [5] Parer JT, King T, Flanders S, et al. Fetal acidemia and electronic fetal heart rate patterns: is there evidence of an association? *J Matern-Fetal Neo M.* 2006;19(5):289–294.
- [6] Cahill AG, Roehl KA, Odibo AO, Macones GA. Association and prediction of neonatal acidemia. *Am J Obstet Gynecol.* 2012;207(3):206.e1–206.e8.
- [7] Clark SL, Hankins GD. Temporal and demographic trends in cerebral palsy-fact and fiction. *Am J Obstet Gynecol.* 2003;188(3):628–633.
- [8] Walsh CA, McMenamin MB, Foley ME, et al. Trends in intrapartum fetal death, 1979–2003. *Am J Obstet Gynecol.* 2008;198(1):47.e1–47.e7.
- [9] Chen HY, Chauhan SP, Ananth CV, et al. Electronic fetal heart rate monitoring and its relationship to neonatal and infant mortality in the United States. *Am J Obstet Gynecol.* 2011;204(6):491.e1–491.e10.
- [10] Ayres-de-Campos D, Spong CY, Chandrachan E. FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography. *Int J Gynaecol Obstet.* 2015;131(1):13–24.
- [11] Rooth G, Huch A, Huch R. Guidelines for the use of fetal monitoring. *Int J Gynaecol Obstet.* 1987;25:159–167.
- [12] SFOG. CTG kort slutversion Sweden: SFOG; 2009. Available from: [https://www.sfog.se/media/17094/ctg\\_kort\\_slutversion.pdf](https://www.sfog.se/media/17094/ctg_kort_slutversion.pdf).
- [13] Amer-Wahlin I, Arulkumaran S, Hagberg H, et al. Fetal electrocardiogram: ST waveform analysis in intrapartum surveillance. *BJOG.* 2007;114(10):1191–1193.
- [14] SFOG. CTG och fosterövervakning Sweden: CTG-utbildning; 2017. Available from: <http://ctgutbildning.se>.
- [15] DSOG. Fosterovervågning under fødslen - indikationer Denmark: DSOG; 2017 [updated 2017 Jan]. Available from: <https://static1.squarespace.com/static/5467abcce4b056d72594db79/t/58bdcef5893fc0f7273b8b7e/1488834299382/Fosteroverv%C3%A5gning+under+f%C3%B8dslen+27.2.docx.pdf>.
- [16] Herbst A, Ingemarsson I. Intermittent versus continuous electronic monitoring in labour: a randomised study. *Br J Obstet Gynaecol.* 1994;101(8):663–668.
- [17] Wiberg N, Kallen K. Fetal scalp blood lactate during second stage of labor: determination of reference values and impact of obstetrical interventions. *J Matern-Fetal Neo M.* 2017;30(5):612–617.
- [18] Bretscher J, Saling E. pH values in the human fetus during labor. *Am J Obstet Gynecol.* 1967;97(7):906–911.
- [19] Yeh P, Emary K, Impey L. The relationship between umbilical cord arterial pH and serious adverse neonatal outcome: analysis of 51,519 consecutive validated samples. *BJOG.* 2012;119(7):824–831.
- [20] Kruger K, Hallberg B, Blennow M, et al. Predictive value of fetal scalp blood lactate concentration and pH as markers of neurologic disability. *Am J Obstet Gynecol.* 1999;181(5):1072–1078.
- [21] McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb).* 2012;22(3):276–282.
- [22] Olofsson P, Noren H, Carlsson A. New FIGO and Swedish intrapartum cardiotocography classification systems incorporated in the fetal ECG ST analysis (STAN) interpretation algorithm: agreements and discrepancies in cardiotocography classification and evaluation of significant ST events. *Acta Obstet Gynecol Scand.* 2018;97(2):219–228.
- [23] Olofsson P, Noren H, Carlsson A, et al. Identifying newborns with umbilical cord blood metabolic acidosis by intrapartum cardiotography combined with fetal ECG ST analysis (STAN): comparison of the new and old FIGO systems to classify cardiotocograms. *J Matern-Fetal Neo M.* 2020;33(3):404–406.
- [24] Marti Gamboa S, Gimenez OR, Mancho JP, et al. Diagnostic accuracy of the FIGO and the 5-tier fetal



- heart rate classification systems in the detection of neonatal acidemia. *Am J Perinatol.* 2017;34(5): 508–514.
- [25] Stephansson O. Graviditetsregistrets årsrapport 2019. Sweden; 2020. Available from: [https://www.medscinet.com/gr/uploads/hemsida/dokumentarkiv/Graviditetsregistrets%20%C3%85rsrapport%202019\\_2.0.pdf](https://www.medscinet.com/gr/uploads/hemsida/dokumentarkiv/Graviditetsregistrets%20%C3%85rsrapport%202019_2.0.pdf)
- [26] Stephansson O, Petersson K, Björk C, et al. The Swedish Pregnancy Register - for quality of care improvement and research. *Acta Obstet Gynecol Scand.* 2018;97(4):466–476.
- [27] Eden RD, Evans MI, Evans SM, et al. The “fetal reserve index”: re-engineering the interpretation and responses to fetal heart rate patterns. *Fetal Diagn Ther.* 2018;43(2):90–104.
- [28] Ayres-de-Campos D, Artero D, Costa-Santos C, et al. Knowledge of adverse neonatal outcome alters clinicians’ interpretation of the intrapartum cardiotocograph. *BJOG.* 2011;118(8):978–984.
- [29] Reif P, Schott S, Boyon C, et al. Does knowledge of fetal outcome influence the interpretation of intrapartum cardiotocography and subsequent clinical management? A multicentre European study. *BJOG.* 2016; 123(13):2208–2217.
- [30] Pinas A, Chandraharan E. Continuous cardiotocography during labour: analysis, classification and management. *Best Pract Res Clin Obstet Gynaecol.* 2016;30:33–47.
- [31] Ugwumadu A. Understanding cardiotocographic patterns associated with intrapartum fetal hypoxia and neurologic injury. *Best Pract Res Clin Obstet Gynaecol.* 2013;27(4):509–536.