Received: 7 September 2022 | Revised: 1 December 2022 | Accepted: 2 December 2022

DOI: 10.1111/aogs.14498

AOGS Acta Obstetricia et Gynecologica Scandinavica

REVIEW

Computerized cardiotocography analysis during labor – A state-of-the-art review

Imane Ben M'Barek[1,2,3] | Grégoire Jauvion[4] | Pierre-François Ceccaldi[2,3,5]

[1]Department of Obstetrics and Gynecology, Assistance Publique Hôpitaux de Paris – Hôpital Beaujon, Clichy La Garenne, France

[2]Université Paris Cité, Paris, France

[3]Health Simulation Department, iLumens, Université Paris Cité, Paris, France

[4]Genos Care, Paris, France

[5]Department of Gynecology-Obstetrics and Reproductive Medicine, Hôpital Foch, Suresnes, France

Correspondence
Imane Ben M'Barek, Service de gynécologie obstétrique, Hôpital Beaujon, 100 boulevard du Général Leclerc, Clichy La Garenne, France.
Email: imane.benmbarek@aphp.fr

Abstract

Cardiotocography is defined as the recording of fetal heart rate and uterine contractions and is widely used during labor as a screening tool to determine fetal wellbeing. The visual interpretation of the cardiotocography signals by the practitioners, following common guidelines, is subject to a high interobserver variability, and the efficiency of cardiotocography monitoring is still debated. Since the 1990s, researchers and practitioners work on designing reliable computer-aided systems to assist practitioners in cardiotocography interpretation during labor. Several systems are integrated in the monitoring devices, mostly based on the guidelines, but they have not clearly demonstrated yet their usefulness. In the last decade, the availability of large clinical databases as well as the emergence of machine learning and deep learning methods in healthcare has led to a surge of studies applying those methods to cardiotocography signals analysis. The state-of-the-art systems perform well to detect fetal hypoxia when evaluated on retrospective cohorts, but several challenges remain to be tackled before they can be used in clinical practice. First, the development and sharing of large, open and anonymized multicentric databases of perinatal and cardiotocography data during labor is required to build more accurate systems. Also, the systems must produce interpretable indicators along with the prediction of the risk of fetal hypoxia in order to be appropriated and trusted by practitioners. Finally, common standards should be built and agreed on to evaluate and compare those systems on retrospective cohorts and to validate their use in clinical practice.

KEYWORDS
cardiotocography deep learning, cardiotocography machine learning, computerized cardiotocography, fetal heart rate monitoring, fetal hypoxia, perinatal morbidity

Abbreviations: CTG, cardiotocography; FHR, fetal heart rate; UC, uterine contractions; ACOG, American College of Obstetricians and Gynecologists; FIGO, International Federation of Gynecology and Obstetrics; AUC, area under the curve; ML, machine learning; DL, deep learning; CNN, convolutional neural network; RCT, randomized controlled trial.

1 | INTRODUCTION

Cardiotocography (CTG) is defined as the recording of fetal heart rate (FHR) and uterine contractions (UC) during pregnancy using an electronic fetal monitor. It is widely used as a screening tool in obstetric practice to determine fetal wellbeing. It is used by practitioners (obstetricians and midwives) during delivery to detect fetal hypoxia, allowing timely intervention before there is permanent damage to the fetus.

CTG monitoring during labor was introduced around 1960, and since then the interpretation of the CTG signals (FHR and UC) has mainly been visually performed by practitioners following common guidelines. The first guidelines were published in 1974 by the American College of Obstetricians and Gynecologists (ACOG),[1] and the International Federation of Gynecology and Obstetrics' (FIGO) guidelines are now used on a widespread basis.[2] CTG signal interpretation is known to be subject to a significant inter- and intraobserver variability.[3–5] More than fifty years later, the benefits of continuous CTG monitoring are still debated and poorly confirmed according to several evaluations comparing it to simpler methods such as intermittent auscultation.[6]

From the 1980s, researchers and practitioners designed computer-aided systems able to automatically process the CTG signals to detect fetal hypoxia during labor. The systems used in clinical practice today are mainly based on features defined in FIGO's guidelines.[7,8] More recently, the publication by the University Hospital in Brno (Czech Republic) of an open-access CTG dataset with 552 clinical cases[9] in 2014, as well as other academic hospitals building larger cohorts,[10] led to a surge of research papers about computerized CTG analysis. Also, the recent emergence of machine learning (ML) and deep learning (DL), and their successful application to similar topics such as computerized analysis of electrocardiograms[11] or electroencephalograms,[12] gave researchers new efficient statistical tools to assist in CTG analysis. The latest systems published, based on ML and DL rather than on the FIGO's guidelines, show promising results when evaluated on retrospective cohorts,[13,14] but many challenges remain to be solved before they can be used in clinical practice.[15]

Here, we present a state-of-the-art review of computerized CTG analysis, with a particular focus on the systems based on the latest ML and DL approaches and released in the very last years. We conclude with our thoughts on the challenges ahead to use those systems in clinical practice.

To select the publications included in this review, we performed a literature search in PubMed, the Cochrane Library, Embase and Google Scholar with the following keywords and MeSH terms: fetal heart rate monitoring, fetal hypoxia, computerized cardiotocography, cardiotocography machine learning, cardiotocography deep learning, cardiotocography randomized controlled trial and cardiotocography interobserver agreement. We selected the papers studying fetal hypoxia and fetal morbidity outcomes. Conference abstracts and editorials were excluded. The list of references in the articles identified were also screened. A full-text analysis of the papers was performed, with a particular focus on the quality of the database (if applicable) and the perceived robustness of the computerized system (Figure 1).

Key message

The use of advanced computerized systems based on the latest machine learning techniques, trained on large databases of cardiotocography data, clinical factors and fetal outcomes, has the potential to successfully assist practitioners in the labor ward and to improve neonatal outcomes.

2 | CTG ANALYSIS BASED ON VISUAL INTERPRETATION

Currently, CTG interpretation during labor is mainly performed visually by practitioners. The first FIGO's guidelines for visual CTG interpretation were introduced in the 1980s, and their latest version, published in 2015, has been summarized by the FIGO Intrapartum Fetal Monitoring Expert Consensus Panel.[2] Those guidelines are based on the following standard features computed on the CTG signals:

- FHR baseline, which is the mean level of the signal evaluated in time periods of 10 min and when the signal is the most stable
- FHR variability, defined as the average amplitude of the signal in one-minute segments
- FHR accelerations (respectively decelerations), which are periods when the signal increases more than 15 bpm above (respectively decreases more than 15 bpm below) the baseline during more than 15 s
- Contractions, defined as increases in the UC signal followed by a symmetric decrease, with a duration between 45 and 120 s

Then, the guidelines define a table enabling the classification of every CTG tracing into three classes (normal, suspicious or pathological) from the value of those standard features. For example, too high or too low FHR baseline values (corresponding respectively to tachycardia and bradycardia), an excessive frequency of contractions (tachysystole), large decelerations lasting more than 3 min or a sinusoidal pattern in FHR are associated with a suspicious or pathological outcome.

The main strength of those guidelines is that they are simple enough to provide practitioners with common rules to interpret the CTG signals in clinical practice during labor. However, we identify two main limitations. First, they miss several features which are known to be linked with fetal hypoxia. For example, maternal and fetal clinical factors such as the gestational age, the fetal estimated weight or the mother's medical history are used by experienced practitioners when interpreting CTG signals, but they are not integrated in the guidelines. Second, several studies where practitioners evaluate the same CTG signals have shown that visual CTG interpretation is

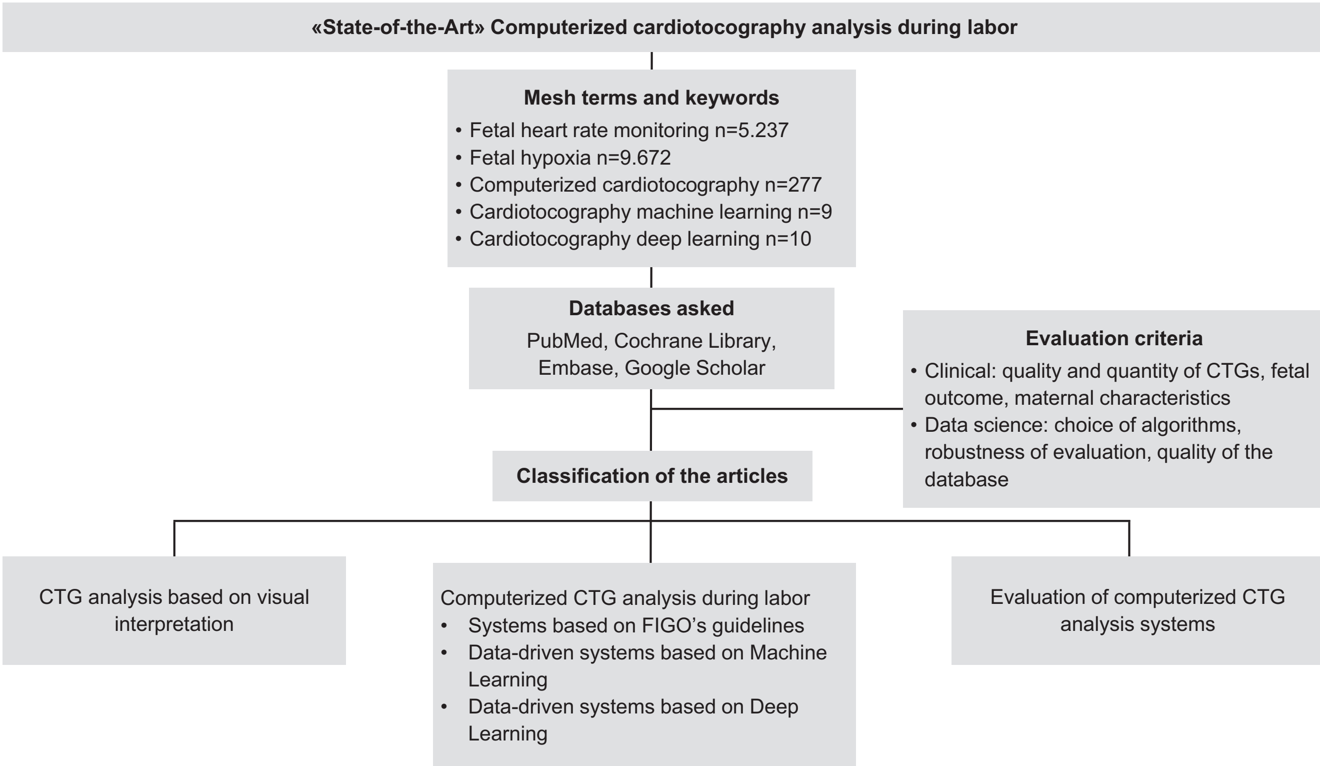

FIGURE 1 Flow chart of the state-of-the-art review of computerized cardiotocography (CTG) analysis.

subject to a significant variability between practitioners, even when they are experienced.[3–5]

Large-scale evaluations of the effectiveness of the current CTG monitoring practice, mainly based on visual interpretation, to detect and prevent fetal hypoxia, have not yet clearly demonstrated that it performs better than other more simple approaches like intermittent auscultation.[6] Those evaluations emphasize the weaknesses of visual interpretation of CTG signals and underline the need for future research on evaluating if computerized interpretation of CTG based on newer algorithms leads to better neonatal and maternal outcomes.

Finally, it is important to note that the guidelines are constantly being challenged. In 2018, Chandraharan et al.[16] introduced a new CTG interpretation methodology that solely relies on physiology-based interpretation for the assessment of fetal wellbeing, as opposed to the FIGO's guidelines. Also, Ayres-de-Campos et al.[17] insist on the need for the guidelines to be simple enough to be used by practitioners, and more objective than they are now.

3 | COMPUTERIZED CTG ANALYSIS DURING LABOR

Although FIGO's guidelines are constantly being challenged, the methods used in clinical practice to interpret CTG have not drastically changed since CTG monitoring was introduced as a screening tool in the 1960s. However, researchers and practitioners worked on designing reliable computerized systems to assist in CTG interpretation since the 1980s, and those efforts have been more important in the last 10 years with the availability of larger clinical databases and the emergence of ML and DL. Table 1 gives a summary of the main computerized systems for CTG analysis which are presented in this review.

3.1 | Systems based on FIGO's guidelines

The first computerized system introduced in the 1980s by Dawes and Redman was built to alert practitioners during pregnancy on the risk of pathological outcome. The system is mostly based on features similar to the ones defined in FIGO's guidelines (baseline, accelerations, decelerations and variability of the signal).[18] Dawes et al. highlighted the advantages of such a computerized analysis.[19] In 2022, Jones et al. published a review of computerized CTG with a focus on the Dawes-Redman system.[20] Then, in the following decades, other computerized systems based on similar methodologies have been built[8,21–23] to be used during labor. The Omniview-SisPorto system (Speculum S.A., Portugal) developed by Ayres-de-Campos et al. consists of a quantitative adaptation of the FIGO's guidelines. The first iteration of the system was published in 1998.[21] An improved version of SisPorto 2.0 was then published in 2000[22] which reports an impressive accuracy (sensitivity 100%, specificity 99%), although it has been evaluated on a small cohort ($n = 85$). SisPorto3.5 was published in 2008[8] and integrates real-time alerts to be used by practitioners, and the latest version of SisPorto 4.0 incorporates FIGO's

TABLE 1 Summary of computerized cardiotocography (CTG) analysis.

	General description	Main systems published
Systems based on FIGO's guidelines (1980s–2010s)	In those systems, simple rules based on FIGO's features (baseline, accelerations, decelerations, variability) are defined to produce relevant indicators of fetal hypoxia for practitioners. Some of those systems have been evaluated on small cohorts.	Dawes-Redman (1980)[19] Omniview-SisPorto (1998)[21] SisPorto 2.0 (2000), SisPorto 4.0 (2017)[22,24] OxSys (2017)[23]
Data-driven systems - machine learning methods applied to features extracted from CTG signals (2010s)	Those systems are based on various features extracted from the CTG signals (including FIGO's features). The risk of fetal hypoxia is predicted from those features using a machine learning model whose parameters are optimized on a CTG database.	Gatellier et al.[31] Abry et al.[32] Spilka et al.[33,34] Czabanski et al.[37]
Data-driven systems - deep learning methods processing raw CTG signals (2020s-now)	In those systems, a deep learning model processes the raw CTG signals, leveraging the latest advances in artificial intelligence techniques. The model is trained on a CTG database, and its accuracy improves with the size of the database.	Fergus et al.[38,39] Ogasawara et al.[41] Mohannad et al.[13] Petrozziello et al.[14,43]

2015 guidelines.[24] Georgieva et al. published the OxSys system in 2017,[23] largely based on one parameter, the decelerative c apacity, which provides an average measure of the downward movements in the FHR signal on a 15 min segment. It is compared with a threshold to trigger real-time alerts for practitioners. On a retrospective database with more than 20.000 recordings, the system is shown to perform better than clinical assessment (as recorded electronically immediately after birth by the attending clinician) to identify moderate and mild fetal distress defined by a pH lower than 7.05 and 7.15, respectively, with a higher sensitivity (36.1% vs 31.0%, $p = 0.06$ and 24.7% vs 22.5%, $p < 0.04$) and a slightly lower false positive rate (14.4% vs 16.3%, $p < 0.001$).

Other computerized systems are used in clinical practice but the methodologies they rely on have not been published and the available descriptions of their characteristics are usually limited, very probably for commercial purpose. Nunes et al. published in 2015 a summary of five systems which are currently integrated with monitoring devices.[7] Those systems all display real-time alerts triggered when a fetal hypoxia suspicion is detected, based on the main features described in the previous section (baseline, accelerations, decelerations and variability). They differ in the thresholds they use to trigger the alerts. The main limitation raised by the authors is that the capacity of those systems to predict adverse neonatal outcomes has not been properly evaluated.

A major technical difficulty faced by researchers to build computerized CTG analysis systems is to automatically estimate the FHR baseline from the raw signal. Although this problem seems relatively simple, a lot of FHR recordings show a high signal variability which makes hard to distinguish between a change in the baseline and an acceleration or a deceleration. Also, the ability for a computerized system to accurately estimate the FHR baseline is key to reach a good accuracy, because the baseline is one of the most important features associated to fetal hypoxia, but also because it is a prerequisite to the identification of accelerations and decelerations. Several research papers have focused on the problem of FHR baseline estimation.[25–30] Dawes et al. used a method based on an analysis of the frequency distribution of the signals.[26,30] Mantel et al. has developed a similar method,[29] which has been compared to the method described in Dawes et al. on 50 recordings: a panel of three experts judged that both methods are equivalent in 36 recordings, and that Mantel et al. is substantially better in the remaining 14 recordings. Lu et al. introduced a somewhat different methodology based on a statistical method named empirical mode decomposition.[28] Ayres-de-Campos et al. compared the baseline estimation performed by SisPorto and by three experienced clinicians on 300 FHR recordings,[25] concluding with an excellent agreement between SisPorto and the clinicians. Finally, Boudet et al. developed a methodology based on a weighted median filter[27] and benchmarked it against 11 published methods. Interestingly, the authors also published a dataset of FHR recordings with the baselines annotated by experts, and it has been used to calibrate the baseline estimation method.

The methods described in this section are mostly based on the computation of standard features defined in FIGO's guidelines, which are compared to arbitrary thresholds triggering a suspicion of fetal hypoxia. In the last years, the availability of large clinical databases of CTG recordings[9,10] with the corresponding fetal outcomes, as well as the emergence of ML and DL, have led to the development of data-driven systems, which are an important focus of the latest research on CTG monitoring.

3.2 | ML-based data-driven systems

A first set of data-driven systems use ML methods to predict fetal hypoxia from features computed on the CTG signals (including the ones defined in FIGO's guidelines).

Gatelier et al. trained a multivariate model on the CTU-UHB dataset and showed that it performs better than visual interpretation based on FIGO's guidelines (AUC = 0.569 for FIGO's guidelines and AUC = 0.719 for the multivariate model).[31] Abry et al. used a ML method to predict fetal hypoxia from a set of 20 features computed from the CTG signals: standard features defined in FIGO's guidelines, but also more advanced spectral and multifractal features.[32] The method is trained and evaluated on the CTU-UHB open dataset[9] and on a dataset collected in Lyon (France), and is based on sparse learning, which consists in selecting a few features while discarding the ones which do not improve significantly the system. The main conclusions are that the system mainly selects the standard FIGO features, confirming their good predictive power, and that a low number of features is enough to build an accurate system. Another conclusion is that the two stages of labor (dilation and active pushing) produce very different FHR dynamics, hence different models must be trained for each stage. Those different conclusions are confirmed in other studies run by Spilka et al.[33,34]

Ribeiro et al. presents a comprehensive review on the most effective statistical methods to build features from the FHR signal,[35] with a particular focus on non-linear features which can then be used in a machine learning model. In another review, Castro et al. focuses on the spectral analysis of FHR signals.[36]

Czabanski et al. proposed a two-step process to analyze FHR signals: in the first step, FIGO's guidelines are applied to identify and discard the cases where a normal outcome is predicted with the highest certainty, and in the second step a ML algorithm is used to refine the detection on the suspicious cases.[37]

3.3 | DL-based data-driven systems

Other data-driven systems use DL and do not require the specification of features built from the CTG signals: they are fed with the raw signals. Applying DL to CTG analysis has the potential to considerably increase the accuracy of fetal hypoxia detection, but also requires larger databases than the other methods introduced previously.

All papers applying DL to CTG analysis use convolutional neural networks (CNNs), a particular form of DL particularly adapted to time-series analysis. Fergus et al. obtain a good accuracy on the CTU-UHB dataset by applying CNNs on fixed-size segments of the CTG signals.[38] In another paper, the same author compared ML and DL methods and showed the superiority of DL methods on the CTU-UHB dataset.[39] Ogasawara et al. applied a CNN to a database with 162 normal cases and 162 abnormal cases, with a decent performance (AUC = 0.73).[40] In another study, Frasch et al. obtained a 94% accuracy by training a CNN on a small database with 36 recordings only.[41] The particularity of this system is that it inputs the images of recordings instead of the raw CTG signals.

Other particularly interesting studies applied DL to much larger clinical databases, which make the evaluations more robust than in the previous studies performed with less than 1.000 records. Mohannad et al. successfully applied a CNN on a clinical database with 38.073 records (AUC = 0.95).[13] Interestingly, the outcome used to train the algorithm is based on the Apgar score, and not on the fetal blood pH like all other papers. Petrozziello et al. applied a CNN on another large database with more than 35.000 births[14,42] and obtained a good accuracy (AUC = 0.81). They also estimate a state-of-the-art accuracy on the open CTU-UHB database (AUC = 0.82), showing that the system generalizes well on the CTG recordings collected in another center. Finally, Zhao et al. applied a CNN on time-frequency images built from the CTG signals and obtained a surprisingly high accuracy on the CTU-UHB dataset (AUC = 0.98).[43]

4 | EVALUATION OF COMPUTERIZED CTG ANALYSIS SYSTEMS

Once they are built and evaluated on retrospective cohorts, a key step is to validate the use of those systems in clinical practice. This is generally performed through Randomized Controlled Trials (RCTs) comparing the system being evaluated with the current practice.

The largest RCT to date has evaluated the INFANT system, a decision-support software to assist practitioners in the interpretation of cardiotocographs, on more than 47.000 deliveries. It concludes that the use of the system does not improve the incidence of poor neonatal outcome, 0.7% in both groups (aRR 1.01, 95% CI: 0.82–1.25).[45] In another RCT performed on about 8.000 deliveries to evaluate a computerized CTG analysis system with real-time alerts for practitioners, Nunes et al. conclude that the system did not significantly reduce the incidence of metabolic acidosis (0.40% vs 0.58%, RR 0.69, 95% CI: 0.36–1.31) or obstetric intervention.[46] A RCT including 720 deliveries reaches an opposite conclusion, showing that reduced risks were observed for all outcomes of interest (hypoxia, cesarean delivery, admission to neonatal intensive care unit) in women monitored using computerized CTG analysis, though the authors note that the small sample size and long recruitment period may overstate the benefits.[47]

Two recent meta-analysis including those three RCTs, representing more than 55.000 patients, have concluded that the use of a computerized CTG system did not impact any of the outcomes evaluated (neonatal acidosis, APGAR scores, mode of delivery and intensive care unit admission) compared to visual interpretation.[47,48]

Finally, several RCTs have been performed to evaluate the STAN system (Neoventa Medical, Sweden), an invasive adjunctive technology which computes the ST analysis based on the fetus's electrocardiogram. Its efficacy to reduce neonatal hypoxia is questionable. According to Norén et al., its use has been associated with a drastic decrease in the cord metabolic acidosis rate in a prospective study run on more than 22.000 pregnancies (0.72% to 0.06%).[49] A meta-analysis by Blix et al. concluded that the use of STAN was associated with a lower rate of metabolic acidosis, but was not associated with a reduction in operative deliveries due to fetal distress.[50] A Cochrane review including more than 27.000 women concluded that the use of STAN made no obvious difference to cesarean section and severe metabolic acidosis.[51]

As far as we know, no system based on ML or DL has yet been evaluated in prospective trials.

5 | DISCUSSION

Although CTG monitoring was introduced more than fifty years ago, it has not yet been demonstrated that it is associated with a significant improvement in neonatal outcomes when compared with simpler approaches such as intermittent auscultation.[6] This may be as a result of several factors, including the biological variability in the ability of the fetus to tolerate hypoxic periods during labor.[53] We are of the opinion that an important factor is the way CTG signals are interpreted visually by practitioners today, a complex process known to induce a significant interobserver variability.[3–5] Using efficient computerized tools to assist practitioners in the interpretation of CTG signals is a very promising way to make the most of CTG monitoring in the labor ward. The tools currently integrated in the monitoring devices are based on relatively simple approaches based on FIGO's guidelines, and they can still be largely improved.

The availability of larger databases[9,10] and of more efficient data analysis tools (eg ML and DL) has led to a surge of research papers introducing computerized CTG analysis systems with a very high accuracy. The most accurate systems as of today are based on DL, and several have been shown to perform better than clinical practice when evaluated on retrospective databases.[14,43] However, there are still challenges ahead which need to be solved before any of those systems can be deployed in clinical practice, and we would like to conclude this review by listing the main challenges and some ideas on how to solve them in the next years.

First, in order to be used in clinical practice, a computerized system must be appropriated and trusted by practitioners. A useful fetal hypoxia prediction system must not only provide practitioners with an accurate prediction of the risk of fetal hypoxia, but it must as well produce interpretable indicators making the prediction understandable. This is quite straightforward for relatively simple systems based for example on FIGO's guidelines that are well known by practitioners, but the more complex becomes the system, the less interpretable it may be. This is a well-known challenge when applying DL methods to healthcare.[54] Also, the use of such systems in clinical practice requires a regular training of the practitioners to make sure they make the most of them.

Another issue we would like to raise is the lack of a common standard to evaluate the systems. A large majority of studies evaluate the systems on retrospective cohorts, and the most popular metric is the AUC. However, it varies a lot between the systems, and some of them reach a surprisingly high performance with an AUC close to 1.[22,43] We encourage the community to build common evaluation methodologies and datasets, to make the reported performances comparable. The CTG challenge developed as part of the Workshop on Signal Processing and Monitoring in Labor[54] is a step in the right direction. The development and sharing of large, open and anonymized multicentric databases of perinatal and CTG data during labor are the conditions for progress in the field.

Beyond their evaluation on retrospective cohorts, the systems must be validated before clinical use in conditions as close as possible to the ones in which they will be used. In particular, the validation methodology must evaluate the whole integration of the computerized system in the labor ward, and not only the piece of software returning a risk of fetal hypoxia. RCTs is an obvious candidate to perform this validation, and it has been applied to several computerized CTG analysis systems with mixed results.[44,49,55] However, given the low incidence of neonatal morbidity (around 0.1%), running such RCTs requires a very large number of patients (at least several dozens of thousands), making them expensive and time consuming. We believe that the community should align on an effective way to validate the effectiveness of computerized CTG analysis systems.

An important specification of data-driven systems is the definition of a pathological outcome. The end-goal of those systems is to reduce the number of adverse neonatal outcomes, and in practice a proxy for this is derived from fetal blood pH or Apgar at 1 or 5 min. The link between fetal acidemia or a low Apgar and an adverse neonatal outcome has been previously studied, for example by Bligard et al.[57] Most studies define a pathological outcome as a pH lower than 7.05 or 7.15, and we found one study that used an outcome based on Apgar at one or 5 min.[13] The specification of the outcome optimized by the system has a key impact on its performance, thus specific studies should be run to align the community on the definition of pathological outcomes.[58]

So far, some teams use invasive adjunctive technologies to detect fetal hypoxia during labor when CTG interpretation is particularly difficult, such as fetal scalp blood sampling and STAN. In addition to their questionable contribution to reducing poor neonatal outcomes,[59,60] these invasive methods are not without risks for the fetus.[61,62] Hence, it is worth improving computerized CTG analysis systems that may replace those invasive technologies at one point.

Finally, it is worth mentioning the important impact computerized CTG analysis could also have outside labor, and especially in remote monitoring during pregnancy. Systems enabling women to monitor FHR and UC at home have been built in the last years, and several studies show that their quality is close to the quality of high-grade cardiotocographs used in hospitals.[62] A qualitative study run by Van den Heuvel et al. concluded that home-based telemonitoring of fetal parameters could be of great use for high risk pregnancies,[63] and another study by Pilarczyk et al. concluded that it should be implemented into everyday obstetric care.[64] As remote CTG monitoring may be performed without any medical practitioner around, the integration of computerized CTG analysis in the process will be particularly useful.

As an overall conclusion, we believe that the use of advanced computerized systems based on ML and DL in CTG monitoring has the potential to bring important benefits by helping practitioners to make better decisions in the labor ward. Several studies have already introduced systems with a very promising performance, but there are challenges ahead that need to be tackled before any use in clinical practice. In particular, those systems need to be trusted and appropriated by practitioners and need to be evaluated with a robust methodology.

FUNDING INFORMATION

Dr Imane Ben M'Barek has university funding for her research on computerized CTG.

ORCID

Imane Ben M'Barek https://orcid.org/0000-0002-2495-0292
Pierre-François Ceccaldi https://orcid.org/0000-0003-4716-9199

REFERENCES

1. American College of Obstetricians and Gynecologists. Fetal heart rate monitoring: guidelines. *ACOG Tech Bull*. 1974;32.
2. 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 Gynaecol Obstet*. 2015;131:13-24.
3. Blackwell SC, Grobman WA, Antoniewicz L, Hutchinson M, Gyamfi BC. Interobserver and intraobserver reliability of the NICHD 3-tier fetal heart rate interpretation system. *Am J Obstet Gynecol*. 2011;205(378):e1-e5.
4. Hruban L, Spilka J, Chudáček V, et al. Agreement on intrapartum cardiotocogram recordings between expert obstetricians. *J Eval Clin Pract*. 2015;21:694-702.
5. Vayssière C, Tsatsaris V, Pirrello O, Cristini C, Arnaud C, Goffinet F. Inter-observer agreement in clinical decision-making for abnormal cardiotocogram (CTG) during labour: a comparison between CTG and CTG plus STAN. *BJOG*. 2009;116:1081-1087. discussion 1087-1088.
6. Alfirevic Z, Gyte GM, Cuthbert A, Devane D. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev*. 2017;2017:CD006066.
7. Nunes I, Ayres-de-Campos D. Computer analysis of foetal monitoring signals. *Best Pract Res Clin Obstet Gynaecol*. 2016;30:68-78.
8. Ayres-de-Campos D, Sousa P, Costa A, Bernardes J. Omniview-SisPorto 3.5 - a central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis. *J Perinat Med*. 2008;36:260-264.
9. Chudáček V, Spilka J, Burša M, et al. Open access intrapartum CTG database. *BMC Pregnancy Childbirth*. 2014;14:16.
10. Houzé de l'Aulnoit A, Parent A, Boudet S, et al. Development of a comprehensive database for research on foetal acidosis. *Eur J Obstet Gynecol Reprod Biol*. 2022;274:40-47.
11. Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol*. 2021;18:465-478.
12. Roy S, Kiral I, Mirmomeni M, et al. Evaluation of artificial intelligence systems for assisting neurologists with fast and accurate annotations of scalp electroencephalography data. *EBioMedicine*. 2021;66:103275.
13. Mohannad A, Shibata C, Miyata K, et al. Predicting high risk birth from real large-scale cardiotocographic data using multi-input convolutional neural networks. *Nonlinear Theory and its Applications, IEICE*. 2021;12:399-411.
14. Petrozziello A, Jordanov I, Aris Papageorghiou T, Christopher Redman WG, Georgieva A. Deep learning for continuous electronic fetal monitoring in labor. *Annu Int Conf IEEE Eng Med Biol Soc*. 2018;2018:5866-5869.
15. O'Sullivan ME, Considine EC, O'Riordan M, Marnane WP, Rennie JM, Boylan GB. Challenges of developing robust AI for intrapartum fetal heart rate monitoring. *Front Artif Intell*. 2021;4:765210.
16. Chandrahan E, Evans S, Krueger D, Pereira S, Skivens S, Zaima A. Physiological CTG Interpretation. 2018. Intrapartum Fetal Monitoring Guideline:33.Guideline. https://physiological-ctg.com/guideline/guideline.html
17. Ayres-de-Campos D, Bernardes J, FIGO Subcommittee. 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.
18. Dawes GS, Moulden M, Redman CW. System 8000: computerized antenatal FHR analysis. *J Perinat Med*. 1991;19:47-51.
19. Dawes GS, Moulden M, Redman CW. The advantages of computerized fetal heart rate analysis. *J Perinat Med*. 1991;19:39-45.
20. Jones GD, Cooke W, Vatish M, Redman C. Computerized analysis of antepartum Cardiotocography: a review. *Maternal-Fetal Medicine*. 2022;4:130-140.
21. Bernardes J, Ayres-de-Campos D, Costa-Pereira A, Pereira-Leite L, Garrido A. Objective computerized fetal heart rate analysis. *Int J Gynaecol Obstet*. 1998;62:141-147.
22. Ayres-de Campos D, Bernardes J, Garrido A, Marques-de-Sá J, Pereira-Leite L. SisPorto 2.0: a program for automated analysis of cardiotocograms. *J Matern Fetal Med*. 2000;9:311-318.
23. Georgieva A, Redman CWG, Papageorghiou AT. Computerized data-driven interpretation of the intrapartum cardiotocogram: a cohort study. *Acta Obstet Gynecol Scand*. 2017;96:883-891.
24. Ayres-de-Campos D, Rei M, Nunes I, Sousa P, Bernardes J. SisPorto 4.0 - computer analysis following the 2015. FIGO guidelines for intrapartum fetal monitoring. *J Matern Fetal Neonatal Med*. 2017;30:62-67.
25. Ayres-de-Campos D, Bernardes J. Comparison of fetal heart rate baseline estimation by SisPorto 2.01 and a consensus of clinicians. *Eur J Obstet Gynecol Reprod Biol*. 2004;117:174-178.
26. Dawes GS, Houghton CR, Redman CW. Baseline in human fetal heart-rate records. *Br J Obstet Gynaecol*. 1982;89:270-275.
27. Boudet S, Houzé de i'Aulnoit A, Demailly R, Peyrodie L, Beuscart R, Houzé de i'Aulnoit D. Fetal heart rate baseline computation with a weighted median filter. *Comput Biol Med*. 2019;114:103468.
28. Lu Y, Wei S. Nonlinear baseline estimation of FHR signal using empirical mode decomposition. In: 2012 IEEE 11th International Conference on Signal Processing. 2012; 1645–1649.
29. Mantel R, van Geijn HP, Caron FJ, Swartjes JM, van Woerden EE, Jongsma HW. Computer analysis of antepartum fetal heart rate: 1. *Baseline Determination Int J Biomed Comput*. 1990;25:261-272.
30. Dawes GS, Visser GH, Goodman JD, Redman CW. Numerical analysis of the human fetal heart rate: the quality of ultrasound records. *Am J Obstet Gynecol*. 1981;141:43-52.
31. Gatellier MA, Jonckheere JD, Storme L, Houfflin-Debarge V, Ghesquière L, Garabedian C. Fetal heart rate variability analysis for neonatal acidosis prediction. *J Clin Monit Comput*. 2021;35:771-777.
32. Abry P, Spilka J, Leonarduzzi R, Chudáček V, Pustelnik N, Doret M. Sparse learning for intrapartum fetal heart rate analysis. *Biomed Phys Eng Express*. 2018;4:034002.
33. Spilka J, Leonarduzzi R, Chudácek V, Abry P, Doret M. Fetal Heart Rate Classification: First vs. Second Stage of Labor. 2016 In 8th International Workshop on Biosignal Interpretation, 2016. [cited 2022 Jul 30].Available from: https://www.semanticscholar.org/paper/Fetal-Heart-Rate-Classification-%3A-First-vs-.-Second-Spilka-Leonarduzzi/bcd13a7980b5e27182d6c1945654b047de99a76d
34. Spilka J, Chudáček V, Koucký M, et al. Using nonlinear features for fetal heart rate classification. *Biomedical Signal Processing and Control*. 2012;7:350-357.
35. Ribeiro M, Monteiro-Santos J, Castro L, et al. Non-linear methods predominant in fetal heart rate analysis: a systematic review. *Front Med (Lausanne)*. 2021;8:661226.
36. Castro L, Loureiro M, Henriques TS, Nunes I. Systematic review of intrapartum fetal heart rate spectral analysis and an application in the detection of fetal Acidemia. *Front Pediatr*. 2021;9:661400.
37. Czabanski R, Jezewski J, Matonia A, Jezewski M. Computerized analysis of fetal heart rate signals as the predictor of neonatal acidemia. *Expert Systems Applications*. 2012;39:11846-11860.
38. Fergus P, Chalmers C, Montanez CC, Reilly D, Lisboa P, Pineles B. Modelling segmented Cardiotocography time-series signals using one-dimensional convolutional neural networks for the early detection of abnormal birth outcomes. *IEEE Trans Emerg Top Computational Intelligence*. 2021;5:882-892.
39. Fergus P, Hussain A, Al-Jumeily D, Huang DS, Bouguila N. Classification of caesarean section and normal vaginal deliveries

using foetal heart rate signals and advanced machine learning algorithms. *Biomed Eng Online*. 2017;16:89.
40. Ogasawara J, Ikenoue S, Yamamoto H, et al. Deep neural network-based classification of cardiotocograms outperformed conventional algorithms. *Sci Rep*. 2021;11:13367.
41. Frasch MG, Strong SB, Nilosek D, Leaverton J, Schifrin BS. Detection of preventable fetal distress during labor from scanned Cardiotocogram tracings using deep learning. *Front Pediatr*. 2021;9:736834.
42. Petrozziello A, Redman CWG, Papageorghiou AT, Jordanov I, Georgieva A. Multimodal convolutional neural networks to detect fetal compromise during labor and delivery. *IEEE Access*. 2019;7:112026-112036.
43. Zhao Z, Deng Y, Zhang Y, Zhang Y, Zhang X, Shao L. DeepFHR: intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network. *BMC Med Inform Decis Mak*. 2019;19:286.
44. INFANT Collaborative Group. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet*. 2017;389:1719-1729.
45. Nunes I, Ayres-de-Campos D, Ugwumadu A, et al. Central fetal monitoring with and without computer analysis: a randomized controlled trial. *Obstet Gynecol*. 2017;129:83-90.
46. Ignatov PN, Lutomski JE. Quantitative cardiotocography to improve fetal assessment during labor: a preliminary randomized controlled trial. *Eur J Obstet Gynecol Reprod Biol*. 2016;205:91-97.
47. Balayla J, Shrem G. Use of artificial intelligence (AI) in the interpretation of intrapartum fetal heart rate (FHR) tracings: a systematic review and meta-analysis. *Arch Gynecol Obstet*. 2019;300:7-14.
48. Campanile M, D'Alessandro P, Della Corte L, et al. Intrapartum cardiotocography with and without computer analysis: a systematic review and meta-analysis of randomized controlled trials. *J Matern Fetal Neonatal Med*. 2020;33:2284-2290.
49. Norén H, Carlsson A. Reduced prevalence of metabolic acidosis at birth: an analysis of established STAN usage in the total population of deliveries in a Swedish district hospital. *Am J Obstet Gynecol*. 2010;202:546.e1-546.e7.
50. Blix E, Brurberg KG, Reierth E, Reinar LM, Øian P. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: a systematic review and meta-analysis of randomized trials. *Acta Obstet Gynecol Scand*. 2016;95:16-27.
51. Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. *Cochrane Database Syst Rev*. 2015;2015(12):CD000116.
52. Johnson GJ, Salmanian B, Denning SG, Belfort MA, Sundgren NC, Clark SL. Relationship between umbilical cord gas values and neonatal outcomes: implications for electronic fetal heart rate monitoring. *Obstet Gynecol*. 2021;138:366-373.
53. McKelvey T, Ahmad M, Teredesai A, Eckert C. Interpretable Machine Learning in Healthcare. BCB '18: Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics. 2018;559–560.
54. Georgieva A, Abry P, Chudáček V, et al. Computer-based intrapartum fetal monitoring and beyond: a review of the 2nd workshop on signal processing and monitoring in labor (October 2017, Oxford, UK). *Acta Obstet Gynecol Scand*. 2019;98:1207-1217.
55. Lopes-Pereira J, Costa A, Ayres-De-Campos D, Costa-Santos C, Amaral J, Bernardes J. Computerized analysis of cardiotocograms and ST signals is associated with significant reductions in hypoxic-ischemic encephalopathy and cesarean delivery: an observational study in 38,466 deliveries. *Am J Obstet Gynecol*. 2019;220:269.e1-269.e8.
56. Bligard KH, Cameo T, McCallum KN, et al. The association of fetal acidemia with adverse neonatal outcomes at time of scheduled cesarean delivery. *Am J Obstet Gynecol*. 2022;227:265.e1-265.e8.
57. Savchenko J, Lindqvist PG, Brismar WS. Comparing apples and oranges? Variation in choice and reporting of short-term perinatal outcomes of term labor: a systematic review of Cochrane reviews. *Eur J Obstet Gynecol Reprod Biol*. 2022;276:1-8.
58. Chandraharan E. Should national guidelines continue to recommend fetal scalp blood sampling during labor? *J Matern Fetal Neonatal Med*. 2016;29:3682-3685.
59. Al Wattar BH, Honess E, Bunnewell S, et al. Effectiveness of intrapartum fetal surveillance to improve maternal and neonatal outcomes: a systematic review and network meta-analysis. *CMAJ*. 2021;193:E468-E477.
60. Sabir H, Stannigel H, Schwarz A, Hoehn T. Perinatal hemorrhagic shock after fetal scalp blood sampling. *Obstet Gynecol*. 2010;115:419-420.
61. Schaap TP, Moormann KA, Becker JH, et al. Cerebrospinal fluid leakage, an uncommon complication of fetal blood sampling: a case report and review of the literature. *Obstet Gynecol Surv*. 2011;66:42-46.
62. Mhajna M, Schwartz N, Levit-Rosen L, et al. Wireless, remote solution for home fetal and maternal heart rate monitoring. *Am J Obstet Gynecol MFM*. 2020;2:100101.
63. van den Heuvel JFM, Teunis CJ, Franx A, Crombag NMTH, Bekker MN. Home-based telemonitoring versus hospital admission in high risk pregnancies: a qualitative study on women's experiences. *BMC Pregnancy Childbirth*. 2020;20:77.
64. Pilarczyk R, Strózik M, Hirnle L. Diagnostic equivalency of mobile CTG devices and remote analysis to conventional on-site nonstress test. *Adv Clin Exp Med*. 2020;29:33-44.

How to cite this article: Ben M'Barek I, Jauvion G, Ceccaldi P-F. Computerized cardiotocography analysis during labor – A state-of-the-art review. *Acta Obstet Gynecol Scand*. 2023;102:130-137. doi:10.1111/aogs.14498