The impact that intense 'fetal-physiology' training has on inter-observer agreement in interpreting intrapartum cardiotocograph traces in comparison to the classical 'pattern-recognition' approach: a prospective observational study

Juan de Dios Gutierrez Henares[1], David Diaz Gomez[1], and Edwin Chandraharan[2]

[1]St George's University Hospitals NHS Foundation Trust
[2]Tianjin Central Hospital of Obstetrics and Gynecology

May 21, 2020

Abstract

Abstract Objective: To explore the level of agreement as well as the contribution of human-factors on CTG interpretation in a hospital where a high intense, 'fetal physiology-based' training on CTG-interpretation was implemented by a dedicated CTG-Team. Design: Prospective observational study Setting: Tertiary Hospital, UK Population: A total of 25 midwives and 7 doctors ([?]10% of staff) Method: interpretation of 5 anonymised colour-printed copies of 5 different CTGs using a questionnaire (160 CTG interpretations) using local CTG-guidelines (pattern-recognition approach) and Types of hypoxia (fetal-physiology). Results: Interpretation of CTG by type of hypoxia compared against CTG-guidelines presented better Proportion of concordance (PC=76.1% vs 61.2%, P=0.006) and slightly better reliability (K= 0.37 (0.35–0.39) vs 0.33 (0.32–0.36)). Doctors rely most in fetal-physiology than midwives who rely most in guidelines. Overall, 68% of the staff felt confident or very confident in CTG interpretation. In general, Self-reported confidence on CTG interpretation and fetal-physiology knowledge increased with the level of seniority. Conclusions: 'Fetal-physiology-based' training increases level of self-reported confidence and level of fetal-physiology knowledge leading to better inter-observer agreement and reliability in CTG interpretation especially when considering type of hypoxia in CTG-classification. Funding: JG collected the data as part of a Self-funded university MSc program. A secondary analysis of the data was performed to elaborate this manuscript. There is no source of funding to declare by the rest of co-authors in this paper. Keywords: Fetal-physiology, Cardiotocography Interpretation, Intrapartum fetal monitoring, Inter-observer agreement,

Tweetable Abstract:

Intense fetal-physiology training by a dedicated CTG-Team helps to improve the levels of confidence and agreement in CTG interpretation

Acronyms used:

- **ACOG:** American Colleague Obstetric and Gynaecologists (USA)
- **Band-5 Midwife:** Midwife within the preceptorship program usually 1-2 years post qualification.
- **Band-6 Midwife:** Experienced or senior midwife who works autonomously, usually 2 to 6 years post qualification

- **Band-7 Midwife:** Very senior midwife, usually <7 years post qualification and acting as co-ordinator or specialist midwife
- **Bpm:** Beats per minute
- **CTG:** Cardiotocograph
- **K:** Kappa-value for reliability
- **NICE:** National Institute for Health and Care Excellence (UK)
- **NICHD:** Eunice Kennedy Shriver National Institute of Child Health and Human Development
- **PC:** Proportion of concordance
- **PIQUARD:** Classificazione di Piquard, 1988 (Italy)
- **STAN:** ST-Analysis (Neoventa Medical(R))

Introduction

Cardiotocograph (CTG) is a screening tool which aims to detect fetal heart responses to an ongoing intrapartum hypoxic or mechanical stress of labour. Nowadays, it is well known that the ability of CTG-monitoring to accurately detect intrapartum hypoxic-stress has been questioned due to its high false-positive rate and the lack of a valid gold-standard for intrapartum-fetal-hypoxia detection to compare with. The recent Cochrane-Systematic-Review on intrapartum-fetal-monitoring[1] concluded that the use of CTG-monitoring increased the rate of C-sections and instrumental deliveries without a significant reduction in the rates of perinatal death or cerebral palsy. Several confidential enquiries into poor perinatal outcomes have highlighted that CTG-misinterpretation is still one of the key avoidable issues. Between 2000 and 2010, the National Health System (NHS) Litigation Authority[2] in the UK identified 300 claims involving CTG-misinterpretation, with an estimated value of £466 million. It is estimated that in the UK, between 500-to-800 babies die or are left with severe brain injuries every year and CTG-misinterpretation has been found to be a contributing factor in 49% of all the cases reported[3]. Misinterpretation of CTGs is mainly subject to two main components, one is the clinical interpretation by the practitioner, and the other, the historical lack of clear consensus by different international and national guidelines. Recently, the use of confusing guidelines on CTG-interpretation based on 'pattern-recognition' have been identified as a source of variability affecting intra and inter-observer agreement[4]. Also, human element is other strong source of variability, as even 'CTG-Experts' have been shown to change their opinion, once they are made aware of the neonatal outcomes[5]. Furthermore, there is still no reliable technology able to alleviate this issue[6]. The only aspect that seems to have had a positive impact in improving inter and intra-observer agreement of CTG-interpretation is intense training and education[7–9]. Nonetheless, there are still controversies about the standardisation and efficacy in the current training schemes offered worldwide[10,11]. Therefore, improving training in CTG-interpretation seems to be crucial to improve perinatal outcomes. Our hypothesis suggests that intense fetal-physiology-based training contribute positively to enhance the inter- and intra-observer agreement as well as levels of self-confidence and knowledge. Although some authors[12], 'appealing to the stone', venture to refute the fetal-physiology-approach in favour to pattern-recognition-approach, it is evident that guidelines based on pattern recognition are contributing to poor perinatal outcomes and increased intrapartum operative interventions[4]. The latest Each Baby Counts Report, published by the Royal College of Obstetricians and Gynaecologists[13], highlights that 33% of cases were due to CTG-Misinterpretation, and in 72% different care may have resulted in different outcomes. In contrast, an intense-training on the use of fetal-physiology to interpret CTG-traces have been reported to be associated with improved perinatal outcomes[9]. The objective of this study was to address the level of agreement, the sources of discrepancies and the associated human factors on CTG-interpretation in staff trained in fetal-physiology-approach from the maternity unit at St George's University Hospitals NHS Foundation Trust in London in order to obtain a deep insight about this *in vogue* method. This hospital, which is one of the largest Teaching Hospitals in London with approximately 5000 births/year, was the first centre in The UK to introduce a mandatory competency testing for all staff providing intrapartum care on CTG-interpretation in 2010 after implemented an intense training in CTG-monitoring based on fetal-physiology-approach, which is provided by a team of highly experienced obstetricians and midwives (CTG-Team in the document). This dedicated CTG-Team

2

has received national awards for its outstanding performance in ensuring a low intrapartum C-section rate and a low hypoxic ischaemic-encephalopathy as compared to other Tertiary Teaching Hospitals in London.

Methods

A total of 25 midwives and 7 doctors, approximately 10% of the total clinical staff, were asked to interpret five anonymised colour-printed-copies of five different CTGs [Fig.1]. Traces were accompanied by the relevant clinical history. Three traces corresponded to ultrasound-transducer recordings, and the other two were CTG-STAN recordings. Along with each copy, a questionnaire with closed and open questions was also provided (supplemental material). The five CTGs were deliberately selected based on the features that give rise to differences in their interpretation. The same questionnaire was also previously filled by the Hospital CTG-Team and was used as theoretical gold-standard for analytical purposes. The questionnaire responses included the categorisation of the CTG-traces, as well as the identification of any ongoing type of hypoxia. In order to classify the traces, local CTG-guidelines (NICE or STAN) were used. Detection of the types of fetal hypoxia on the questionnaire was based on the described criteria in the scientific literature[14,15]: gradually evolving hypoxia, subacute hypoxia, acute hypoxia, and chronic hypoxia. The questionnaire also allowed the quantification of several aspects: (1) the proportion of concordance, between the CTG-Team and clinical staff, in CTG-classification by 'CTG-guidelines' as well as by identification of 'types-of-hypoxia', (2) the inter-rater (inter-observer) reliability within the staff, (3) the background knowledge in CTG-interpretation and (4) the level of self-reported confidence.

Statistical analysis

The CTG categorisations provided by the CTG-Team compared with the categorisation given by the staff was assessed by proportion of concordance (PC) with 95% confidence interval (CI). The staff inter-observer reliability was assessed by Fleiss-Kappa value (K). K-values were interpreted according to Landis and Koch[16] recommendations: a K<0.20 was considered poor, 0.20-0.40 slight, 0.41-0.60 fair, 0.61-0.80 substantial, and 0.81-1.00 almost perfect. The rest of the proportions that were mainly descriptive were expressed as raw percentages without CI. Comparison of different PC was assessed by chi-square test with a significant level set a P<0.001. Comparison of K-values was assessed following Cumming and Finch[17] where K were considered non-significantly different if the 95% CI overlaps. The statistical analysis was generated using the Real-Statistics Resource-Pack software (Release-4.3) for Excel-Microsoft-Office 2015 and IBM SPSS-Statistics for Windows, Version-25.0. Armonk, NY:IBM Corp.2017.

Ethical approval

Data was obtained as part of a university MSc-program and therefore followed the ethical guidelines of the UK universities in addition of the permission of the Hospital Local Ethics Committee and the voluntary participation of the staff. No patient identifiable data were used in the study.

Results

CTG interpretation: Categorisation and types of hypoxia.

In total, 160 CTG full interpretations, five for each participant were examined. The analysis of the differences between the CTG-Team and the clinical staff on CTG-interpretation applying local CTG-guidelines are displayed in table-1. Overall, the categorisation of CTG using the correspondent local-guideline presented a PC (95% CI) = **61.2%**(53.6%–68.8%), representing a moderate agreement against the CTG-Team and a K (95% CI) = **0.33** (0.316–0.362), representing a fair reliability. However, if the CTGs are being interpreted by types of hypoxia, the PC= **76.1%** (69.4%–82.8%) and K=**0.37**(0.35–0.39). Consequently, the identification of type hypoxia compared against local-guidelines as method of CTG interpretation presented better PC

(76.1% vs 61.2%, P=0.006) and slightly better reliability (K 0.37 (0.35–0.39) vs 0.33 (0.32–0.36)). In comparison with other methods of interpretation based in pattern-recognition and published under peer-review; interpretations by types-of-hypoxia present the higher proportion of agreement, and also, better reliability than studies with similar sample of observers [Table-2]

Background knowledge

The staff were asked to rank from 1-5 which source of knowledge helped them most in analysing each CTG. The options given were: 1) uses of current guidelines, 2) own knowledge in fetal-physiology, 3) previous experience, 4) opinion of someone more senior and 5) similar case(s) previously discussed during a CTG meeting/training. Midwives reported that the background knowledge on which they rely the most are guidelines first (25.8%), and fetal-physiology second (22.2%) followed by experience (20.4%), discussion in previous CTG-meetings (16.3%) and opinion (15.2%). Doctors relied mostly on fetal-physiology (28.9%), experience (20.5%) and meeting (20.5%) were ranked both in second position with same percentage followed by guidelines (19.3%), and opinion (10.7%) [Table-3; Fig.2-4]

Self-reported level of confidence

The staff were asked to rank the level of confidence over 7 possible points from 'not confident at all' to 'very confident'. Overall, 68% of them feel confident or very confident with CTG interpretation. Within the midwifery group, the most confident or very confident were Band-7 midwives (94.7%) followed by Band-6 (64.5%) and Band-5 (41.7%). Doctors followed a similar pattern to midwives. The most confident to very confident were the Consultants (100%), followed by senior doctors (90%) and junior doctors (57.1%). [Table-4;Fig:5]

Discussion

Since the purported rationale of having different categorisation of a CTG-trace is to identify the risk of the potential hypoxia, our study shows that it is more practical to directly state whether a fetus is exposed to a hypoxic stress and the type of ongoing fetal hypoxia, if any. This may help avoid the use of confusing terminology such as 'intermediate' 'suspicious' or 'pathological' CTG-traces, which have no correlation with neonatal outcomes[18]. Also, it is worth to mention that our method to calculate PC imply a double agreement: first between staff and second against the gold-standard. Therefore, we suggest this method enhance the validity of our agreement results.

Sources of discrepancy: Pattern-recognition vs. Fetal-Physiology

The staff that did not agree with the diagnosis of the **CTG-1**and described it as suspicious or pathological were led by the number of uterine contractions shown in the tocograph and not by non-reassuring features on the cardiograph. This suggests that features that are not formally part of the CTG-guideline table may interfere with the overall interpretation. The intense fetal-physiology training ensures that the trained staff is also able to consider any ongoing excessive uterine activity contributing to abnormal features on the CTG-trace. Although being vigilant for any deviation from normality is crucial in maternity services, clinicians should also bear in mind that an over-diagnosis may be equally harmful, as it may lead to expediting the delivery of a healthy fetus. An interesting data for discrepancy was noted in **CTG-2**, where up to 10 different nomenclatures were used to describe decelerations. Although, none of those categories and nomenclature would lead to different management other than imminent delivery, the use of appropriate terminology stipulated by the guidelines was not followed. This reflects the inherent flaws in any guideline which is based on 'pattern-recognition' which relies on the morphological classification of decelerations, as this would lead to significant inter and intra-observer variability. According to the CTG-Team, the **CTG-3** baseline is 108bpm, was a non-reassuring feature as stipulated by the guidelines and thus, the CTG must be categorised as suspicious. However, the staff who categorised the CTG as normal did so because they considered that the base line was [?]110bpm. The problem that arises from this 2bpm difference is that a

4

base line of 108bpm in a term baby can be perfectly normal, but it can be (strictly speaking) categorised as a non-reassuring feature. Consequently, if any other non-reassuring feature appears while the baseline is defined as suspicious, the CTG would be categorised as pathological. A similar scenario was seen on the **CTG-4** as the main discrepancy was categorising the trace as intermediary (under STAN-guidelines) due to a base line of >150bpm. Understanding the importance of accurately (and physiologically) interpreting baseline is crucial to avoid over diagnosis leading to potential unnecessary interventions because incorrect assessment would lead to incorrect management.**CTG-5** only presented one complicated-deceleration with a reassuring baseline. However, in the context of STAN-guidelines, which differentiated between different types of decelerations, but do not specify the number of decelerations required per determined period of time, promotes confusion in the CTG-categorisation. Similar to CTG-2, the confusion arise from naming the decelerations or mixing the guidelines producing a confusion that can reduce the rate of agreement only on the basis of the terminology. This highlights the role played by some guidelines based on 'pattern-recognition' in promoting confusion amongst clinicians.

Source of knowledge

When midwives progress from Band-5 to Band-6, logically, they start relying more in their own experience and less on the opinion of someone more senior. Most importantly, the data show that the more senior the midwife, the more reliance on the fetal-physiology to interpret the CTG-traces and a diminished reliance on the CTG-guideline until they become Band-7. This last group reported experience as the least valued option and their decisions are mostly based on the use of guidelines followed the by knowledge of fetal-physiology. A possible explanation to this phenomenon amongst Band-7 (labour co-ordinators) may be their crucial role in having an 'overall' responsibility which could create conflicts between taking defensive decisions following a closed written-guideline or trusting the fetal-physiology. Similar scenario as Band-7 is seen on senior doctors, but not in Consultants. However, since doctors also increase reliance on the understating of the fetal-physiology along seniority, it is likely that the CTG intense-teaching is promoting a switch from pattern-recognition to a physiological-approach amongst staff, this can be easily visualised in the radial graphs provided [Fig. 2-4].

Confidence on CTG-interpretation

The level of confidence varies according to professional grade. Both, midwives and doctors gain self-confidence as they progress in their respective careers. Band-7 midwives (i.e. labour ward co-ordinators) reported a higher level of confidence than junior and senior doctors. This is likely due to the intense 'cascade training' on CTG-interpretation provided to Band-7 midwives by the CTG-Team to ensure that the unit is always staffed by a co-ordinator with an excellent knowledge of fetal-physiology. The lower proportions of being confident or very confident are among Band-5 midwives (i.e. newly appointed or junior midwives). This is understandable, considering that they are the professionals who are most likely to seek a senior opinion. In contrast, 100% of consultants felt confident or very confident. However, it is also interesting to highlight a considerable disagreement in CTG-interpretation between the consultants who took part in this study. This is likely due to the incorporation of individual 'experience', disregarding the guidelines or fetal-physiology by some consultants. Therefore, it is important to appreciate that some degree of overconfidence and/or non-concordance may exist amongst senior clinicians in any maternity team due to their experience. Therefore, a multidisciplinary-team approach to CTG-interpretation by improving the knowledge of fetal-physiology may help improve concordance and reliability in CTG-interpretation.

Importance of Fetal-Physiology training and multi-professional approach

Our study highlights the challenges that arise when pattern-recognition is in place. On one hand, relying mostly on CTG-guidelines, especially in junior staff, could be seen as a "horse-blinder" producing inability to see and understand a wider clinical picture such as an appropriate fetal-heart-rate base-line, ongoing chorioamnionitis, maternal pyrexia, meconium stained liquor, etc. This is usually manifested by lower level of self-confidence in CTG-interpretation. On the other hand, in the more senior staff, there is a chance

of taking a more defensive and interventionist approach by relying more in CTG-guidelines and 'personal experience' than in the actual physiology and clinical picture. This could be manifested by an overconfidence status. Therefore, it is vital to ensure all staff receive intense training on fetal-physiology and the types of intrapartum-hypoxia, so the pattern-recognition approach do not trump physiological and scientific principles underpinning intrapartum fetal heart rate monitoring[18]. To support the above, our study demonstrates that intense training on fetal-physiology not only improve K and PC but also increase knowledge and self-confidence in CTG-interpretation. This will contribute to reduce the variation in the management of labour, and hopefully, will improve intrapartum maternal and perinatal outcomes. Additionally, instead of using multiple CTG-guidelines based on pattern-recognition with confusing terminologies and different 'features', we suggest the use of 'types of intrapartum-hypoxia' to classify CTG-traces as a default method. This will contribute to delineate better the fetal ability to respond and compensate to an hypoxic insult, which is the corner-stone of intrapartum-CTG. Similar findings were reported in a recent study[19] which analysed 52,187 births over an 11-years period, which reported 81% agreement between clinicians when 'types of hypoxia' were used to classify the CTG-Trace, instead of using guidelines based on 'pattern-recognition'.

Strengths and limitations

To our best knowledge, this is the first study which analysed inter-observer variability amongst 32 mid-wives and obstetricians of different grades and experience who have undergone an intense training on fetal-physiology. Secondly, we had a dedicated CTG-Team, who have expertise on CTG-interpretation as they have published extensively in this area, and conduct CTG-Masterclasses in approximately 14 countries every year, who were used as the 'gold-standard'. Thirdly, in addition to inter-observer variability, we also analysed subjective levels of confidence on CTG-interpretation. The main limitation was the restriction to a single centre. However, the authors felt that it was best to conduct this study in a centre where had received an intense training on fetal-physiology, and a mandatory competency testing on CTG-interpretation. Secondly, it may be argued that clinicians were provided with only 20-minutes of the CTG-trace instead of the whole trace, we accept that 'Cycling'[8,20], one the most important CTG-features, could not be evaluated properly. However, to determine the type of hypoxia, it was felt that a 20-minute trace was sufficient and it reflected the real life situation, where clinicians are expected to make crucial decisions based on short segments of the CTG-trace. Thirdly, the authors accept some may argue that number of observer was small. However, this was a complex study assessing inter-observer variability, and although only 32 clinicians took part in the study, a total 160 CTG-traces were analysed. Many studies on inter-observer variability on CTG traces have used less than 10 clinicians[21–29].

Conclusions

This paper demonstrates that continuous education and an intense 'fetal physiology-based' CTG-training by a specialised CTG-team increase knowledge in fetal-physiology and produces higher levels of staff confidence reflecting better levels of agreement and reliability. Classification of CTG-Traces by 'type of intrapartum-hypoxia' is preferable to CTG-guidelines. However, if CTG-guidelines are sine-*qua-non* element of the maternity unit, they should be simple and easy to use, and these should be backed up by immediate availability of senior input with appropriate knowledge of fetal physiology able to recognise any ongoing hypoxic process. This approach may help reduce the pitfalls of pattern recognition amongst more junior members of staff. Development of a specialized "CTG-Team" formed by consultants and midwives to educate staff, and to review and discuss CTG-traces and outcomes may help to create a multidisciplinary approach resulting in inter-observer variability reduction and increased staff confidence in CTG interpretation.

Conflict of interest:

No conflict of interest has been declared by the authors.

Contribution to authorship:

JG conceived the study and performed data collection. JG, DD and EC undertook data analysis and interpreted results. All authors contributed to the writing of the document and approved the final manuscript.

Details of ethical approval:

Data was obtained as part of a St. Georges University of London MSc-program (LA: HP7203/4/XY approved and signed on 22.06.2017) in addition of the permission of the Local NHS trust and the voluntary participation of the staff. No patient identifiable data were used in the study.

Reference List

1. Alfirevic Z, Devane D, Gyte GML, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour.Cochrane Database Syst Rev 2017; 2017. doi: 10.1002/14651858.CD006066.pub3

2. Anderson A. Ten years of maternity claims: an analysis of the NHS Litigation Authority data – key findings.Clin Risk 2013; 19: 24–31.

3. Royal College of Obstetricians and Gynaecologists. Each Baby Counts Clinical Engagement Forum – Results Slidepack. 2017

4. Chandraharan E. Intrapartum care: An urgent need to question historical practices and 'non-evidence'-based, illogical foetal monitoring guidelines to avoid patient harm.J Patient Saf Risk Manag 2019; 24: 210–7.

5. Ayres-de-Campos D, Arteiro D, Costa-Santos C, Bernardes J. Knowledge of adverse neonatal outcome alters clinicians' interpretation of the intrapartum cardiotocograph.BJOG An Int J Obstet Gynaecol 2011; 118: 978–84.

6. Steer PJ, Kovar I, McKenzie C, Griffin M, Linsell L. Computerised analysis of intrapartum fetal heart rate patterns and adverse outcomes in the INFANT trial.BJOG 2019; 126: 1354–61.

7. Gyllencreutz E, Hulthen Varli I, Lindqvist PG, Holzmann M. Reliability in cardiotocography interpretation – impact of extended on-site education in addition to web-based learning: an observational study.Acta Obstet Gynecol Scand 2017; 96: 496–502.

8. Chandraharan E. Handbook of CTG Interpretation. Cambridge: Cambridge University Press, 2017. .

7

9. Chandraharan E, Lowe V, Ugwumadu A, Arulkumaran S. Impact of fetal ECG (STAN) and competency based training on intrapartum interventions and perinatal outcomes at a teaching hospital in London: 5 year analysis.BJOG 2013; 120: 428.

10. Ugwumadu A, Steer P, Parer B, Carbone B, Vayssiere C, Maso G, Arulkumaran S. Time to optimise and enforce training in interpretation of intrapartum cardiotocograph.BJOG An Int J Obstet Gynaecol 2016; 123: 866–9.

11. Pehrson C, Sorensen JL, Amer-Wahlin I. Evaluation and impact of cardiotocography training programmes: A systematic review.BJOG An Int J Obstet Gynaecol 2011; 118: 926–35.

12. Sholapurkar SL. Myths at the core of Intrapartum Cardiotocography Interpretation – Risks of false Ideology, Prospect theory and way forward.Clin Obstet Gynecol Reprod Med 2019; 5: 1–9.

13. Royal College of Obstetricians and Gynaecologists. Each Baby Counts: 2019 Progress Report.London: RCOG; 2020; 1–58.

14. Ugwumadu A. Understanding cardiotocographic patterns associated with intrapartum fetal hypoxia and neurologic injury.Best Pract Res Clin Obstet Gynaecol 2013; 27: 509–36.

15. Chandraharan E. Physiological CTG - Guideline [Internet]. [cited 2020 May 2]. Available from: https://physiological-ctg.com/guideline.html

16. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data.Biometrics 1977; 33: 159.

17. Cumming G, Finch S. Inference by eye confidence intervals and how to read pictures of data.Am Psychol 2005; 60: 170–80.

18. Pinas A, Chandraharan E. Continuous cardiotocography during labour: Analysis, classification and management.Best Pract Res Clin Obstet Gynaecol 2016; 30: 33–47.

19. Yatham SS, Whelehan V, Archer A, Chandraharan E. Types of intrapartum hypoxia on the cardiotocograph (CTG): do they have any relationship with the type of brain injury in the MRI scan in term babies?J Obstet Gynaecol 2019; 1–6.

20. Pereira S, Chandraharan E. Recognition of chronic hypoxia and pre-existing foetal injury on the cardiotocograph (CTG): Urgent need to think beyond the guidelines.Porto Biomed J 2017; 2: 124–9.

21. Ojala K, Makikallio K, Haapsamo M, Ijas H, Tekay A. Interobserver agreement in the assessment of intrapartum automated fetal electrocardiography in singleton pregnancies.Acta Obstet Gynecol Scand 2008; 87: 536–40.

22. Chauhan SP, Klauser CK, Woodring TC, Sanderson M, Magann EF, Morrison JC. Intrapartum nonreassuring fetal heart rate tracing and prediction of adverse outcomes: interobserver variability.Am J Obstet Gynecol 2008; 199: 623.e1-5.

23. Figueras F, Albela S, Bonino S, Palacio M, Barrau E, Hernandez S, Casellas C, Coll O, Cararach V. Visual analysis of antepartum fetal heart rate tracings: inter- and intra-observer agreement and impact of knowledge of neonatal outcome.J Perinat Med 2005; 33: 241–5.

24. Bernardes J, Costa-Pereira A, Ayres-de-Campos D, van Geijn HP, Pereira-Leite L. Evaluation of interobserver agreement of cardiotocograms.Int J Gynaecol Obstet 1997; 57: 33–7.

25. Lotgering FK, Wallenburg HC, Schouten HJ. Interobserver and intraobserver variation in the assessment of antepartum cardiotocograms.Am J Obstet Gynecol 1982; 144: 701–5.

26. Blackwell SC, Grobman WA, Antoniewicz L, Hutchinson M, Bannerman CG. Interobserver and intraobserver reliability of the NICHD 3-Tier Fetal Heart Rate Interpretation System.Am J Obstet Gynecol 2011; 205: 378.e1-378.e5.

8

27. Rei M, Tavares S, Pinto P, Machado AP, Monteiro S, Costa A, Costa-Santos C, Bernardes J, Ayres-De-Campos D. Interobserver agreement in CTG interpretation using the 2015 FIGO guidelines for intrapartum fetal monitoring.Eur J Obstet Gynecol Reprod Biol 2016; 205: 27–31.

28. Ghi T, Morganelli G, Bellussi F, Rucci P, Giorgetta F, Rizzo N, Frusca T, Pilu G. Cardiotocographic findings in the second stage of labor among fetuses delivered with acidemia: a comparison of two classification systems.Eur J Obstet Gynecol Reprod Biol 2016; 203: 297–302.

29. Santo S, Ayres-de-Campos D, Costa-Santos C, Schnettler W, Ugwumadu A, Da Graca LM, Clode N, Pinto L, Centeno M, Carvalho R, Rocha P, Reis I, Afonso M, Castro C, Aserkoff R, Golen T, Ricciotti H, Vadnais M, Scott J, Rana S, Quant H, Caldwell L, Chandraharan E, Bhide A, Lack N, Narang L, Muhammad S, Devarajan S, Squires R, Affors K, Ugwumadu L. Agreement and accuracy using the FIGO, ACOG and NICE cardiotocography interpretation guidelines.Acta Obstet Gynecol Scand 2017; 96: 166–75.

Tables

CTG number	CTG -Team	CTG -Team	% of Clinical Staff that gave correct	% of Clinical Staff tha
	Categorisation	Type of Hypoxia	category	type of hypoxia
CTG-1	Normal	Non-Hypoxia	78%	91%
CTG-2	Abnormal	Sub-Acute	69%	59%
CTG-3	Suspicious	Non-Hypoxia	13%	84%
CTG-4	Normal	Non-Hypoxia	97%	100%
CTG-5	Normal	Non-Hypoxia	50%	47%

Table 1 Proportion of concordance between staff and the CTG-Team for each individual CTG trace on giving the correct categorisation as well as the correct type of hypoxia.

	N_o	N_t	N_i	Method of CTG interpretation	Proportion of agreement (95% CI)	Intra-observer reliability (95% CI)
Blackwell et al. (2011)[26]	3	120	360	NICHD	Not reported	0.45 (not reported)
Ghi et al (2016)[28]	2	246	492	RCOG	0.72 (Not reported)	0.58 (0.50-0.66)
				PIQUARD	0.60 (Not reported)	0.50 (0.44 – 0.56)
Rei et. al (2016)[27]	6	151	906	FIGO	0.60(0.56 – 0.64)	0.39 (0.33 – 0.45)
Santo et al (2017)[29]	27	151	4077	FIGO	0.64 (0.61 – 0.67)	0.37 (0.31 – 0.43)
				ACOG	0.73 (0.70 – 0.76)	0.15 (0.10 – 0.21)
				NICE	0.55 (0.51 – 0.58)	0.33 (0.28 – 0.39)
Fetal-Physiology-trainning	32	5	160	Local guidelines (NICE / STAN)	0.61 (0.53 – 0.68)	0.33 (0.32 – 0.36)
				Type of hypoxia	0.76 (0.69 – 0.82)	0.37 (0.35 – 0.39)

Table 2 Comparison of the proportion of concordance and reliability in studies with similar methodology and published under peer-review

	Guidelines (%)	Fetal Physiology (%)	Experience (%)	Opinion (%)	CTG-Meeting (%)
Midwives-5	27.3	18.7	16.0	20.2	17.8
Midwives-6	24.8	22.4	21.9	14.0	16.9
Midwives-7	34.0	31.5	2.5	16.8	15.1
All Midwives	**25.8**	**22.2**	**20.4**	**15.2**	**16.3**
Junior Drs	15.3	26.5	16.3	11.6	30.2
Senior Drs	29.2	24.0	29.2	11.7	5.8
Consultants	13.3	33.3	26.7	6.7	20.0
All Doctors	**19.3**	**28.9**	**20.5**	**10.7**	**20.5**

Table 3 distribution of the percentages of the staff source knowledge to face the interpretation of the cardiotocography trace. The table shows the staff divided by role (Doctor or midwives) and subdivide by grading within the role. Band-5, Midwife within the preceptorship program usually 1-2 years post qualification. Band-6, experience midwives who works autonomously, usually 2 to 6 years post qualification. Band-7 midwife very senior midwife, usually <7 years post qualification and acting as co-ordinator or specialist midwife. Junior doctor usually 3 years post qualification. Senior doctor usually 4-8 years post qualification. Consultant >9-10 post qualification.

Level of confidence	Band-5 (%)	Band-6 (%)	Band-7 (%)	Junior Drs (%)	Senior Drs (%)	Consultants (%)	Staff (%)
very confident or Confident	41.7	64.5	94.7	57.1	90.0	100.0	68.0
Somewhat confident	45.8	31.6	5.3	42.9	10.0	0.0	28.1
neither confident nor Unconfident	4.2	1.3	0.0	0.0	0.0	0.0	1.3
Somewhat Unconfident	0.0	1.3	0.0	0.0	0.0	0.0	0.7
Unconfident or Not confident at all	8.3	1.3	0.0	0.0	0.0	0.0	2.0

Table 4 Table 3 distribution of the percentages of the level of self-reported staff confidence who took part in the study. The table shows the staff divided by role (Doctor or midwives) and subdivide by grading within the role. Band-5, Midwife within the preceptorship program usually 1-2 years post qualification. Band-6, experience midwives who works autonomously, usually 2 to 6 years post qualification. Band-7 midwife very senior midwife, usually <7 years post qualification and acting as co-ordinator or specialist midwife. Junior doctor usually 3 years post qualification. Senior doctor usually 4-8 years post qualification. Consultant >9-10 post qualification. Midwife wit
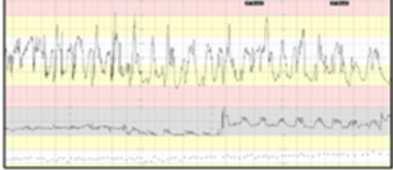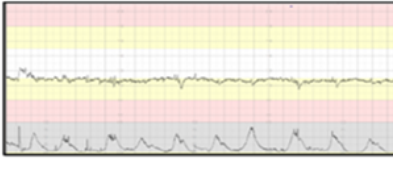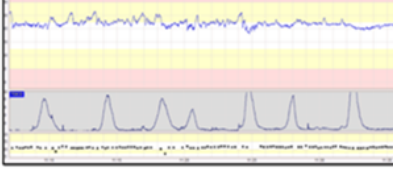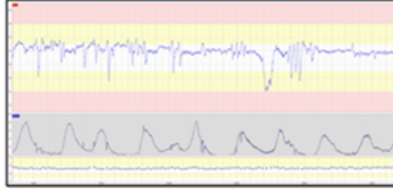
Figures

Figure-1: Copy of the five anonymised colour-printed CTG traced with their relevant clinical history that has been used in the study.

Figure-2: Radial graph that represents the background knowledge distribution of all doctors and all midwives. The staff were asked to rank from 1-5 which source of knowledge helped them most in analysing each CTG. The options given were: 1) uses of current guidelines, 2) own knowledge in fetal physiology, 3) previous experience, 4) opinion of someone more senior and 5) similar case(s) previously discussed during a CTG meeting/training.

Figure-3: Radial graph that represents the background knowledge distribution of all midwives clustered by midwifery banding. The staff were asked to rank from 1-5 which source of knowledge helped them most in analysing each CTG. The options given were: 1) uses of current guidelines, 2) own knowledge in fetal physiology, 3) previous experience, 4) opinion of someone more senior and 5) similar case(s) previously discussed during a CTG meeting/training.

Figure-4: Radial graph that represents the background knowledge distribution of all doctors clustered medical grading. The staff were asked to rank from 1-5 which source of knowledge helped them most in analysing each CTG. The options given were: 1) uses of current guidelines, 2) own knowledge in fetal physiology, 3) previous experience, 4) opinion of someone more senior and 5) similar case(s) previously discussed during a CTG meeting/training

Figure-5: Pie-Chart illustrating the proportional distribution of the level of confidence of the total staff taking part in the study. The staff were asked to rank the level of confidence over 7 possible points from 'not confident at all' to 'very confident'. The point were clustered post-hoc for analytical purposes.

N	CTG	Clinical history given
1		**G2 P1 38+3 Spontaneous onset** Previous SVD Epidural No STAN® analysis available, Fetal Scalp electrode monitoring **6cm Dilated** Temp 36.8 HR 96pm BP 120/81
2		**G1 P0 39+4** 2X ST EVENT: T/QRS rise 0.06 IOL for GDM Fully Dilated **Pushing** Temp 36.2 HR 78bpm BP139/90
3		**G3 P2 41+5** Previous C/S (FTP) 3 years ago SVD 2 years ago **5CM dilated** Temp 36.2 HR 73bpm BP 118/63
4		**G1 P0 41+2** Spontaneous onset High uterine arteries at 21wk (normal growth and dopplers after) SROM at 2am clear, 4cm dilated at 6:30 **6cm at 10:35am, Current time 11:35am** No changes in baseline sin admission No previous CTG available HR 85bpm BP 126/73 Temp 36.8
5		**G2 P0** (previous miscarriage 11wks) **39+5 weeks** IOL for GDM porpess® x1 ARM 16h ago at 3cm clear liquor Oxytocin infusion started 10h ago O/E: 8cm LOA at spines at 05:25 (current time 6am) Temp 36.9 HR 79bpm BP 126/69

Doctors and Midwives Background knowledge distribution

Midwives background knowledge distribution

13

Doctors background Knowledge distribution

Level of confidence Total Staff