The evaluation of a tabular application of the NICE guidelines for universal interpretation of non-stress test (NST) and cardiotocograph(CTG)

Dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, as partial fulfillment for the MMed degree (Obstetrics and Gynaecology)

11/11/2014

Dr Noxolo B. Jack

P O Box 4012

Halfway House

0832765720

noxolob@yahoo.co.uk

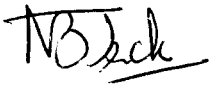
Supervisors:

1. **Dr Poovangela Naidoo**
   Senior specialist, CHB
   Academic hospital.
2. **Prof E. Buchmann,**
   District Clinical
   Specialist, Johannesburg.

DECLARATION

I, Dr Noxolo Brenda Jack, declare that this dissertation is my own work. It is being submitted to the Faculty of Health Sciences, University of Witwatersrand, as partial fulfillment for the MMed degree in Obstetrics and Gynaecology.

Dr N.B. Jack



Date: 11/11/2014

ACKNOWLEDGEMENTS

I would like to extend my warmest gratitude to Dr Poovangela Naidoo for her remarkable assistance and guidance in this study; Prof Eckhart J Buchmann for assisting with data analysis and guidance; the midwives and doctors who participated in this study for their time and enthusiasm; Prof F Guidozzi who assisted with the compilation and development of this document; and my husband for his continuous support in making sure the work is complete.

Table of contents

DECLARATION 1
ACKNOWLEDGEMENTS 2
ABSTRACT 5
1. INTRODUCTION 7
1.1 Electronic fetal heart rate monitoring (EFM) 7
1.2 The use of electronic fetal monitoring 8
1.3 The efficacy of electronic fetal monitoring 8
1.4 The interpretation of electronic fetal heart rate components 9
1.5 Assessment of electronic fetal heart rate patterns 13
1.6 Guidelines for the interpretation of NSTs and CTGs 14
1.6.1The International Federation of Gynecology and Obstetrics (FIGO) guidelines 15
1.6.2 The National Institute of Child Health and Human Development (NICHHD)/ American College of Obstetricians and Gynecologists (ACOG) guidelines 18
1.6.3 The NICE/Royal College of Obstetricians and Gynaecologists (RCOG) guidelines 21
1.7 The NICE guidelines for the use of EFM 21
1.7.1 Interpretation of EFM according to the National Institute of Health and Clinical Excellence (NICE) guidelines 22
1.8 Problem Statement 26
2. OBJECTIVES 26
3. METHODS 27
3.1 Setting 27
3.2 Study design and study population 28
3.3 Data Collection 28
3.3.1 The 2-part questionnaire 28
3.3.2 First Questionnaire 28
3.3.3 Second Questionnaire 30
3.4 Data analysis 35
3.5 Ethics 36
4. RESULTS 37
5. DISCUSSION 44
5.1 Medical rank of participants 44

5.2 The non-systematic assessment of traces 45
5.3 Systematic evaluation of traces using the NICE guidelines tabular model with a scoring modification 45
5.3.1 Interpretation of fetal heart rate components 45
5.3.2 Classification of fetal heart rate patterns using the NICE guidelines 47
5.3.3 Problem areas identified in the interpretation of NSTs and CTGs 47
5.3.4 Limitations of this study 48
6. CONCLUSION 49
7. REFERENCES 50
APPENDIX A First questionnaire 53
APPENDIX B Second questionnaire 54
APPENDIX C Ethics approval 55
..... 55

ABSTRACT

Objective: To assess consensus in the interpretation of cardiotocographs (CTGs) and non-stress tests (NSTs) between different grades of obstetric clinical staff by comparing assessment of traces by non-systematic eyeballing with assessment of traces using a tabular approach suggested by the National Institute of Health and Clinical Excellence (NICE) guidelines for interpretation of CTGs and NSTs, and to identify components of NSTs and CTGs where medical personnel experience difficulty with interpretation.

Design: Prospective observational study.

Setting: Maternity units of the tertiary care hospitals for the teaching and training of the Witwatersrand University postgraduates, interns and midwives.

Participants: Midwives, advanced midwives, interns, medical officers, registrars and specialists working in the above-mentioned maternity units.

Method: Participants were recruited at the time of formal gatherings and departmental meetings in the various institutions. Each participant was given five traces that were a combination of NSTs and CTGs to interpret and assess in a non-systematic way using three categories: baby well; baby requires further surveillance; and baby needs immediate delivery. The same participants were then given the same set of traces in a different sequence for interpreting in a systematic way using the tabular approach from the NICE guidelines on electronic fetal monitoring with a scoring modification.

Main outcome measure: Differences in interpretation of CTGs by different grades of staff, and degree of certainty between study participants in the different assessment systems.

Results: Twenty seven specialists, 25 registrars, 21 medical officers, 10 interns and 15 midwives participated. There were varying interpretations by individuals in both the non-systematic assessment and the systematic assessment using the NICE tabular application, with best agreement in Trace 3 (77% and 84% respectively). In the non-systematic assessment, there was a statistically significant difference in the assessment of traces 1, 2 and 4 between the different grades of staff (P-values<0.01, 0.03 and <0.01 respectively). There was no statistically significant difference when the traces were assessed using the NICE guidelines tabular application (P-values; Trace1 >0.99, Trace2=0.27, Trace 3 = 0.76, Trace 4 = 0.15 and Trace5 = 0.35).Certainty of the evaluation by the participants was determined if 75% or more of the participants agreed on a classification. Using the NICE guidelines, there was uncertainty (failure to agree on classification by 75% or more of the participants) with baseline variability, accelerations, decelerations and overall assessment of the CTG in most of the traces.

Conclusion: There is no uniformity in the assessment of traces by midwives, interns, medical officers, registrars and specialists. Some uniformity in the interpretation of traces and reduction in inter-observer variation is attained by the use of the NICE guidelines tabular application. However, baseline variability, accelerations, and decelerations remain a problem in the interpretation of NSTs and CTGs using the NICE guidelines.

1. INTRODUCTION

The ability to diagnose fetal life through auscultation of the fetal heart marked the beginning of fetal monitoring. An interest as to how to recognize changes in fetal heart rate in order to prevent perinatal mortality mounted. A considerable advancement in the techniques of auscultation led to the development of electronic fetal monitoring.[1] Over the years the techniques of electronic fetal monitoring improved but the interpretation of fetal heart rate (FHR) patterns has not been optimized and hence the tool for determining fetal well-being has only been as good as the interpreter. The challenge is no longer to perfect the test but to unify and hence improve interpretation.

1.1 Electronic fetal heart rate monitoring (EFM)

Electronic fetal monitoring (EFM) employs the use of ultrasound to evaluate the fetal heart rate in-utero and hence fetal brain function. There are two categories of monitoring:

1) Non-Stress test (NST) measures the FHR by means of an ultrasonic sensor in the absence of uterine contractions where uterine contractions are perceived as the stress.

2) Cardiotocography (CTG) measures the FHR by means of an ultrasonic sensor, in relation to uterine contractions which are monitored using a tocodynamometer.

Interpretation of an NST relies on the assessment of the following parameters:

1) Baseline FHR
2) Baseline FHR variability
3) The presence of accelerations
4) Interpretation of decelerations

The interpretation of a CTG includes assessment of the above-mentioned components of an NST and in addition the assessment of uterine activity represented on the tocograph.

Effective EFM requires a correctly performed test, adequate interpretation of the results, and an appropriate response based on the interpretation.[1, 2]

1.2 The use of electronic fetal monitoring

Electronic fetal monitoring was intended to detect fetal hypoxia before the development of fetal acidosis and ensuing perinatal mortality, and thus improve the birth outcome and prevent cerebral palsy.[3, 4, 5] This technique became widely used from the 1960s, and is a widely accepted method for fetal surveillance in pregnancy and during labour.[1, 6,7]

1.3 The efficacy of electronic fetal monitoring

Many studies have been conducted to assess the efficacy of EFM. In 2001 a Cochrane meta-analysis of EFM showed only a minor beneficial effect on the incidence of neonatal seizures

(RR, 0.5; 95% CI, 0.31-0.8), no effect on the incidence of cerebral palsy (RR, 1.74; 95% CI, 0.9 -3.11) or perinatal death (RR, 0.85; 95% CI, 0.59–1.23), but an increase in operative vaginal delivery (RR, 1.16; 95% CI, 1.0 –1.32) and caesarian section (RR, 1.66; 95% CI,1.3–2.13).[8] The negative predictive value of EFM ranges from 93% - 99% .[2] The positive predictive value is much lower and ranges between 8% and 26%.[2] This indicates a high false positive and a low false negative rate for fetal compromise.

The 4th Confidential Enquiry into Stillbirths and Deaths in Infancy in the United Kingdom highlighted problems related to the use and interpretation of continuous tocographic (CTG) traces where CTG misinterpretations as well as inappropriate action in the presence of abnormal fetal heart rate changes are the most common alleged negligences.[9, 10, 11]

1.4 The interpretation of electronic fetal heart rate components

In practice, a CTG trace is usually interpreted as a whole based on the contribution of the four features described earlier. Depending on the system used for CTG interpretation, the FHR components are defined and further categorized. After categorizing the individual components, the CTG trace is classified. A complete clinical understanding of EFM is necessary for one to appreciate changes in the above-mentioned characteristics.

Variations in the interpretation of FHR components and patterns between observers, even amongst experts, have led to extensive research on the topic. Most studies have demonstrated that identification of FHR baseline among observers was fair to good, whereas interpretation or identification of accelerations and decelerations vary .[1, 2, 12, 13]

Baseline variability has been found to be the most difficult to interpret visually.[14] Several studies revealed no parity in the classification of FHR variability.[1,5,12,15] It has been noted that clinicians have a tendency to over-report abnormalities especially with reduced variability and variable decelerations.[14]

In 1978, Trimbos et al. were the first to investigate inter- and intra-observer agreement in the interpretation of CTG tracings.[13] Concerns were raised regarding the reliability of visual interpretation of CTG tracings soon after EFM was introduced. Five observers were given 100 CTGs to interpret independently on two occasions.[13] They found the highest level of agreement among observers in the assessment of baseline FHR, followed by decelerations and the lowest level of agreement in the recognition of accelerations.[13] Ninety seven percent of traces were scored the same by all five observers in terms of baseline FHR description, whereas 65% in the detection of decelerations and only 34% in the recognition of accelerations.[13]

In 1982, Lotgering et al. found that the overall intra-observer agreement was high ($k = 0.7 - 0.8$) whereas the inter-observer agreement was low ($k = 0.09 - 0.69$) for all variables assessed.[15] The highest level of agreement was in the assessment of baseline fetal heart rate for both intra-observers ($k = 0.83 - 0.89$) and inter-observers ($k = 0.53 - 0.69$). Agreement was lowest in the assessment of baseline variability for both intra-observers ($k = 0.12 - 0.46$) and inter-observers ($k = 0.09 - 0.32$).[15]

Nielsen et al. also noted considerable intra- and inter-observer variability in 1987 when he gave four obstetricians 50 tracings on two occasions to identify pathological tracings.[5] Twenty one percent of the traces were interpreted differently at the second observation. Only 22% of the traces were assessed the same by all four obstetricians.

In 1993, Donker et al. asked 21 experienced obstetricians to describe and classify13 CTG tracings.[2] They were then given clinical information on each patient and were asked to assess the fetal condition and to propose obstetric management. Results showed fair agreement among observers on the classification of baseline FHR, accelerations and decelerations (overall k = 0.48, range 0.04 – 0.53). There was poor agreement for baseline variability (k = 0.16, range 0.01 – 0.21) and for the type of deceleration. The overall CTG interpretation, clinical assessment and obstetric management also showed poor agreement among the observers. It was concluded that there is still use of ambiguous terminology and definitions in the assessment of fetal heart rate tracings. Standardization of terminology, definitions and criteria for FHR analysis were suggested as one of the solutions. The use of computerized analysis was thought to be another helpful tool as it would assess fetal heart rate patterns consistently.[2]

In 1996, a comparison between two experienced obstetricians, two non-experienced obstetricians, and a computer system on reproducibility of CTG readings was done by Todros et al.[16] They observed that the overall reproducibility among the observers was fair to good (k = 0.05 – 0.67) for most variables. There was poor agreement (k = 0.10 – 0.48) between the observers and computer readings for baseline FHR. The agreement was fair to good for baseline variability (k = 0.16 – 0.74), acceleration (k = 0.37 – 0.64) and decelerations (k = 0.41 – 0.54) between the observers and the computer readings. The authors concluded that the use of a computer system should overcome the problem of intra- and inter-observer variability.[16]

In 1997, Bernades et al. evaluated inter-observer agreement in the classification of CTGs by giving three experienced obstetricians 33 CTGs to interpret using the FIGO guidelines and relevant clinical information.[12] Agreement was fair to good in the assessment of baseline, accelerations and uterine contractions. There was poor agreement in the detection of

decelerations. Agreement in the assessment of baseline variability was good if it was considered normal, but poor if considered abnormal. It was observed that CTG reproducibility is still poor even when performed by experienced clinicians with consensual guidelines and access to clinical information. It was also noted that reproducibility was good in detection of more gross and stable CTG events but poor in detection of subtle alterations.[12] These findings were attributed to ambiguous definitions, interdependence of definitions, difficulty in eyeball evaluation of subtle CTG alterations and difficulty in systematic and disciplined assessment of CTGs by busy clinicians. It was suggested that guidelines for EFM should be revised and more precise definitions of CTG events included. A more disciplined method of CTG analysis by the clinicians and the use of computerized analysis were also suggested.[12]

The National Institute of Child Health and Human Development in the United States conducted a Research Planning Workshop in 1997 on electronic fetal heart rate monitoring.[17] The purpose of the workshop was to propose standardized and unambiguous definitions for fetal heart rate tracings. Recommendations for the interpretation of fetal heart rate patterns were laid down.[17]

In the assessment of midwives in 2005 Devane et al. asked 28 midwives to interpret three intrapartum CTGs independently on two separate occasions.[5] The midwives assessed the individual heart rate components and classified the tracings into normal, suspicious and pathological using the FIGO and Family Health International guidelines. They found that the overall intra-observer agreement in CTG interpretation was fair to good to excellent ($k = 0.48 – 0.92$), similar to Lotgering et al. findings in 1982.[15] The overall inter-observer agreement in the interpretation of CTG tracings was in the upper limit of fair to good category ($k = 0.65 – 0.74$). Inconsistent with the literature, the agreement was highest in the classification of decelerations ($k = 0.79$).[5] Agreement was the lowest in the assessment of baseline variability ($k = 0.5$), a finding

which is in keeping with the literature. The agreement was highest in suspicious tracings (k = 0.77) and lowest in normal tracings (k = 0.54), a finding which is in contrast with the literature. Devane et al. attributed the small number of tracings used in the study as the possible reasons for the conflicting results.[5]

1.5 Assessment of electronic fetal heart rate patterns

In one study it was found that errors of interpretation are reduced if FHR traces are categorized as a whole, with reference to individual features.[18, 19] The four features of the fetal heart rate must be described (baseline FHR, baseline variability, accelerations and decelerations). Each feature is then categorized into normal, suspicious or abnormal. Based on the contribution of all the features, the whole CTG is classified. The interpretation of CTGs was more consistent in normal traces than seen with suspicious or pathological traces.[18, 19, 20] In 1999, Ayres-de-Campos et al. used the same setting as Bernades et al. Observers classified the CTGs as normal, suspicious, or pathological according to the FIGO guidelines.[20] They were also requested to decide on one of the clinical management options: no action, close monitoring or immediate intervention. The overall agreement in the classification of FHR tracings was in the lower limit of fair to good category (k = 0.48; 95% CI 0.34-0.62). There was reasonable agreement for normal tracings (Pa = 0.62; 95% CI 0.51-0.73) and poor agreement for suspicious (Pa = 0.42; 95% CI 0.34-0.50) and pathological tracings (Pa = 0.25; 95% CI 0.14-0.36). Agreement was significantly better for no action (Pa = 0.79; 95% CI 0.68-0.89) than for close monitoring (Pa = 0.14; 95% CI 0.02-0.43) or immediate intervention (Pa = 0.38; 95% CI 0.21-0.56). It was noted that disagreement occurred mainly when situations diverged from normal.[20]

Similar to previous studies were the findings by Blix et al. in 2003. They evaluated inter-observer variation in the assessment of 845 labour admission CTGs by midwives and obstetricians in the clinical setting and two experts in the non-clinical setting. The traces were assessed according to the Ingemarsson and Ingemarsson classification as reactive, equivocal or ominous (Table 1). Proportions of agreement were high for reactive traces and low for equivocal and ominous traces.[18]

Based on the above findings, it seems that inter-observer and intra-observer consistency of interpretation remains poor even with the introduction of specific guidelines for interpreting fetal heart rate traces.[6]

Table 1 Classification of fetal heart rate patterns according to Ingemarsson et al.[18]

Reactive	Two accelerations (more than 15 beats, more than 15 seconds) in 10 – 20 minutes. Traces without accelerations but normal baseline and variability.
Equivocal	Absence of accelerations with reduced variability or silent pattern, but normal frequency. Uncomplicated (normal variability) tachycardia or bradycardia. Uncomplicated variable decelerations.
Ominous	Silent pattern and tachycardia or bradycardia. Late decelerations, complicated variable decelerations. Heart frequency less than 100 beats per minute or prolonged deceleration.

1.6 Guidelines for the interpretation of NSTs and CTGs

Several organizations have produced fetal monitoring guidelines with partly assimilated concepts, some contradictory aspects, some with excessively complex rules.[1] The three-class classification system approach that classifies CTGs into one of three categories (normal, abnormal or indeterminate) has now been accepted by all because of its simplicity.[21]

1.6.1The International Federation of Gynecology and Obstetrics (FIGO) guidelines

The first guidelines for the use of fetal monitoring were produced in a workshop that was organized by the International Federation of Gynecology and Obstetrics (FIGO) Subcommittee on Standards in Perinatal Medicine in 1985.[22] The aim of these guidelines was to assist in the correct use of electronic fetal heart rate monitoring. A consensus was reached on several aspects of fetal heart rate monitoring including the method, terminology, indications, technique and interpretation. The FIGO guidelines apply separate classification criteria for antepartum and intrapartum fetal heart rate patterns.[22] The guidelines are shown in Tables 2 and 3.

Table 2 Definitions of FHR features according to the FIGO guidelines[22]

TERM	DEFINITION
Baseline fetal heart rate	The mean level of the fetal heart rate when this is stable, accelerations and decelerations being absent. It is determined over a time period of 5 or 10 min and expressed in beats/min (bpm).
Baseline variability	Oscillations of fetal heart rate around its mean level (long term variability). It is usually only quantitated by description of the amplitude of the oscillations around the baseline heart rate. Under physiological conditions the fetal beat-to-beat intervals are constantly subject to small changes called short term variability. These cannot be reliably interpreted by the naked eye using the standard equipment.
Accelerations	Transient increase in heart rate of 15 beats/min or more and lasting 15 seconds or more.
Decelerations	Transient episodes of slowing of fetal heart rate below the baseline level of more than 15 beats/min and lasting 10 seconds or more.
Sinusoidal pattern	Regular cyclic changes in the fetal heart rate baseline, such as the sine wave. The characteristics of the pattern being: the frequency is less than 6 cycles/min, the amplitude is at least 10 beats/min and the duration should be 20 min or longer.

***FIGO guidelines do not provide a definition of early, variable or late decelerations.**

Table 3 Classification of fetal heart rate patterns according to the FIGO guidelines.[22]

ANTEPARTUM	INTRAPARTUM
Normal	**Normal**
- Baseline heart rate between 110 and 150bpm	- Baseline heart rate between 110 and150bpm
- Amplitude of heart rate variability between 5 and 25 bpm	- Amplitudeof heart rate variability between 5 and 25 bpm
- Absence of decelerations except for sporadic, mild decelerations of very short duration	
- Presence of two or more accelerations during a 10-minute period	
Suspicious (if any of the following signs are present)	**Suspicious (if any of the following are present)**
- Baseline heart rate between 150 and 170 bpm or between 100 and 110 bpm	- Baseline heart rate between 150 and 170 or between 100 and 110 bpm
- Amplitude of variability between 5 and 10 bpm for more than 40minutes	- Amplitude of variability between 5 and 10 bpm for more than 40minutes
- Increased variability above 25 bpm	- Increased variability above 25 bpm
- Absence of accelerations for more than 40 minutes	
- Sporadic decelerations of any type unless severe	- Variable decelerations
Pathological (when any of the following signs are present)	**Pathological**
- Baseline heart rate below 100 or above 170 bpm	- Baseline heart rate below 100 or above 170 bpm
- Persistence of heart rate variability of less than 5 bpm for more than 40 minutes minutes	- Persistence of heart rate variability of less than 5 bpm for more than 40
- Periodically recurring and repeated decelerations of any type	- Severe variable decelerations or severe repetitive early decelerations
- Sporadic and non-recurrent decelerations of the following types: severe variable, prolonged, late	- Prolonged decelerations - Late decelerations: the ominous trace is a steady baseline without baseline variability and with small decelerations after each contraction
- A sinusoidal pattern	- A sinusoidal pattern

1.6.2 The National Institute of Child Health and Human Development (NICHHD)/ American College of Obstetricians and Gynecologists (ACOG) guidelines

The National Institute of Child Health and Human Development Research Planning Workshop held in 1997 proposed standardized definitions for fetal heart rate patterns which were then reviewed and pattern interpretation updated in 2008. The 2008 workshop was conducted in partnership with the American College of Obstetricians and Gynecologists (ACOG) and the Society of Maternal Fetal Medicine. The NICHHD/ACOG guidelines apply to intrapartum patterns, but can also apply to antepartum patterns.[19] The guidelines are shown in Tables 4 and 5.

Table 4 Definitions of fetal heart rate features according to NICHHD/ ACOG[19]

Baseline	The mean FHR rounded to increments of 5 beats per minute (bpm) during a 10 minute segment, excluding: - periodic or episodic changes - periods of marked FHR variability - segments of baseline that differ by more than 25 bpm The baseline must be for a minimum of 2 mins in any 10 - minute segment, or the baseline for that time period is indeterminate. In this case one may refer to the prior 10-minute window for determination of baseline. -Normal FHR baseline: 110-160 bpm -Tachycardia : FHR baseline >160 bpm -Bradycardia: FHR baseline < 110 bpm
Baseline variability	Fluctuations in the baseline FHR that are irregular in amplitude and frequency. Variability is visually quantitated as the amplitude of peak-to-trough in bpm. - absent – amplitude range undetectable - minimal – amplitude range detectable but 5 bpm or fewer - moderate (normal) – amplitude range 6-25bpm - marked – amplitude range greater than 25 bpm
Acceleration	A visually apparent abrupt increase (onset to peak in less than 30 seconds) in the FHR. At ≥32weeks of gestation, an acceleration has a peak of ≥15 bpm above baseline, with duration of ≥15seconds but < 2 minutes from onset to return. Before 32 weeks of gestation, an acceleration has a peak of ≥10 bpm above baseline, with a duration of ≥ 10 seconds but< 2 minutes from onset to return.

	Prolonged acceleration lasts ≥ 2 minutes but < 10 minutes. If an acceleration last ≥ 10 minutes, it is a baseline change.
Early deceleration	Visually apparent usually symmetrical gradual decrease and and return of the FHR associated with a uterine contraction. A gradual FHR decrease is defined as from the onset to the FHR nadir of ≥30 seconds. The decrease from the FHR is calculated from the onset to the nadir of the deceleration. The nadir of the deceleration occurs at the same time as the peak of the contraction. In most cases the onset, nadir and recovery of the deceleration are coincident with the beginning, peak and ending of the contraction, respectively.
Late deceleration	Visually apparent usually symmetrical gradual decrease and and return of FHR associated with a uterine contraction. A gradual FHR decrease is defined as from the onset to the FHR nadir of ≥30 seconds. The decrease from the FHR is calculated from the onset to nadir of the deceleration. The deceleration is delayed in timing, with the nadir of the deceleration occurring after the peak of the contraction. In most cases, the onset, nadir and recovery of the deceleration occur after the beginning, peak, or ending of the contraction, respectively.
Variable deceleration	Visually apparent abrupt decrease in FHR. An abrupt FHR decrease is defined as from the onset of the deceleration to the beginning of FHR nadir of < 30 seconds The decrease from the FHR is calculated from the onset to nadir of the deceleration. The decrease in FHR is ≥ 15bpm lasting ≥ 15 seconds and < 2 minutes in duration. When variable decelerations are associated with uterine contractions, their onset, depth, and duration commonly vary with successive uterine contractions.
Prolonged deceleration	Visually apparent decrease in FHR below the baseline. Decrease in FHR from the baseline that is ≥ 15bpm lasting ≥ 2 minutes but < 10 minutes in duration. If a deceleration lasts ≥ 10 minutes, it is a baseline change.
Sinusoidal pattern	Visually apparent, smooth, sine wave-like undulating pattern in FHR baseline with a cycle frequency of 3-5 per minute which persists for 20 minutes or more.

Table 5 Classification of fetal heart rate pattern according to NICCHD/ ACOG guidelines, 2008.[23]
(Reproduced from Reviews in Obstetrics and Gynecology)[23]

Category I

Normal tracings, which are strongly predictive of normal fetal acid-base status at the time of observation and can be followed in a routine manner without any specific action required, include *all* of the following:
• Baseline rate: 110-160 beats/min
• Moderate variability
• Absence of any late or variable decelerations
• Early decelerations may or may not be present
• Accelerations may or may not be present

Category II
Indeterminate tracings, although not predictive of abnormal fetal acid-base status, cannot be classified as Category I or III and thus require evaluation and continued surveillance and reevaluation. These tracings are not infrequently encountered in clinical care, and include any of the following:
• Baseline rate
- Tachycardia
- Bradycardia not accompanied by absent baseline variability
• Baseline FHR variability
- Minimal baseline variability
- Absent baseline variability not accompanied by recurrent decelerations
- Marked baseline variability
• Absence of induced accelerations after fetal stimulation (e.g. scalp stimulation, vibroacoustic stimulation, direct fetal scalp sampling, transabdominal halogen light)
• Periodic or episodic decelerations
- Recurrent variable decelerations accompanied by minimal or moderate baseline variability
- Prolonged deceleration more or equal to 2 min but less than 10 min
- Recurrent late decelerations with moderate baseline variability
- Variable decelerations with other characteristics, such as slow return to baseline, "overshoots," or "shoulders"

Category III
Abnormal tracings, which are predictive of abnormal fetal acid-base status at the time of observation, require prompt evaluation and initiation of expeditious attempts to resolve the abnormal FHR pattern, such as provision of maternal oxygen, change in maternal position, discontinuation of labor stimulation, treatment of maternal hypotension, or additional efforts. These tracings include *either*:
• Absent baseline FHR variability along with *any* of the following:
- Recurrent late decelerations
- Recurrent variable decelerations
- Bradycardia
• Sinusoidal pattern

1.6.3 The NICE/Royal College of Obstetricians and Gynaecologists (RCOG) guidelines

The most recent guidelines on fetal monitoring were developed by the National Institute of Health and Clinical Excellence (NICE) and produced by the Royal College of Obstetricians and Gynecologists (RCOG) in 2001. These will be elaborated in detail below as this study is on the NICE guidelines. The NICE guidelines were chosen for evaluation for these reasons: the fetal heart rate features are described (as with the other guidelines), the FHR features are further classified into reassuring, non-reassuring and abnormal (not done by the other two guidelines, FIGO and NICHHD/ACOG). This makes the overall assessment of the tracing simpler as one does not have to remember each and every point on the three-class classification system. One just needs to score how many features fall into each of the categories and make an assessment.

From a recent study by Ayres-de-Campos et al. comparing the FIGO guidelines, the RCOG/NICE guidelines, and the ACOG/NICHHD guidelines; it became clear that the lack of consensus in many aspects of universally accepted guidelines will remain a major limitation to the effectiveness of CTG as a fetal monitoring technique.[21] Further simplification of the existing guidelines and generalized assimilation of concepts and recommendations has been suggested as an important step to address this difficulty.[21]

1.7 The NICE guidelines for the use of EFM

In 2001 the National Institute of Health and Clinical Excellence (NICE) in the UK with the Department of Health developed clinical guidelines for the use of electronic fetal monitoring. They defined EFM as the use of electronic fetal heart rate monitoring for the evaluation of fetal

wellbeing in labour. The aim of these guidelines was to develop criteria for the use of EFM, indications for use, definitions of normal and abnormal parameters, and to evaluate methods for improving interpretation of CTG.[1]

1.7.1 Interpretation of EFM according to the National Institute of Health and Clinical Excellence (NICE) guidelines

Interpretation of EFM requires recognition of a normal pattern, and the results outside the normal range increase the probability of fetal compromise.[9]

Table 6 NICE Guidelines for definitions and descriptions of individual features of fetal heart rate traces[1]

TERM	DESCRIPTION
Baseline fetal heart rate	The mean level of the FHR when this is stable, excluding accelerations and decelerations. It is determined over a period of 5 or 10 minutes and expressed in beats per minute (bpm). Preterm foetuses tend to have values towards the upper end of this range. A trend to a progressive rise in the baseline is important as well as the absolute values.
Normal Baseline FHR	110–160 pm
Moderate bradycardia	100–109 bpm
Moderate tachycardia	161–180 bpm
Abnormal bradycardia	<100 bpm
Abnormal tachycardia	>180 bpm
Baseline variability	The minor fluctuations in baseline FHR occurring at three to five cycles per minute. It is measured by estimating the difference in beats per minute between the highest peak and lowest trough of fluctuation in a one-minute segment of the trace
Normal baseline variability	≥ 5 bpm between contractions
Non-reassuring baseline variability	< 5 bpm for ≥40 minutes but < 90 minutes
Abnormal baseline variability	<5 bpm for ≥90 minutes

Accelerations	Transient increases in FHR of ≥15 bpm and lasting ≥15seconds. The significance of no accelerations on an otherwise normal CTG is unclear
Decelerations	Transient episodes of slowing of FHR below the baseline level of > 15 bpm and lasting ≥15 seconds
Early decelerations	Uniform, repetitive, periodic slowing of FHR with onset early in the contraction and return to baseline at the end of the contraction
Late decelerations	Uniform, repetitive, periodic slowing of FHR with onset mid to end of the contraction and nadir >20 seconds after the peak of the contraction and ending after the contraction. In the presence of a non-accelerative trace with baseline variability < 5 bpm, the definition would include decelerations < 15 bpm
Variable decelerations	Variable, intermittent periodic slowing of FHR with rapid onset and recovery. Time relationships with contraction cycle are variable and they may occur in isolation. Sometimes they resemble other types of deceleration patterns in timing and shape
Atypical variable decelerations	Variable decelerations with any of the following additional components: i. Loss of primary or secondary rise in baseline rate, ii. Slow return to baseline FHR after the end of the contraction. iii. Prolonged secondary rise in baseline rate, iv. Biphasic deceleration, v. Loss of variability during deceleration, vi. Continuation of baseline rate at lower level.
Prolonged deceleration	An abrupt decrease in FHR to levels below the baseline that lasts at least 60–90 seconds. These decelerations become pathological if they cross two contractions, i.e. >3 minutes
Sinusoidal pattern	A regular oscillation of the baseline long-term variability resembling a sine wave. This smooth, undulating pattern, lasting at least 10minutes, has a relatively fixed period of 3–5 cycles per minute and amplitude of 5–15 bpm above and below the baseline. Baseline variability is absent. These ranges of baseline are not associated with hypoxia in the presence of accelerations, with normal baseline variability and no decelerations

Table 7. Classification of fetal heart rate features according to the NICE Guidelines[1]

Feature	Baseline rate (bpm)	Variability	Decelerations	Acceleration
Reassuring	110 – 160	≥ 5	None	Present
Non-reassuring	100 – 109 161 – 180	<5 for >40 minutes but <90 minutes	Early decelerations Variable decelerations Single prolonged deceleration < 3 minutes	The absence of accelerations with an otherwise normal CTG is of uncertain significance.
Abnormal	<100>180 Sinusoidal pattern for more than 10 minutes	< 5 for ≥ 90 minutes	Atypical variable decelerations Late decelerations Single prolonged deceleration > 3 minutes	

Table 8. Classification of CTG according to the NICE Guidelines[1]

Category	Definition
Normal	A CTG where all four features fall into the 'reassuring' category
Suspicious	A CTG where one of the features falls into 'non-reassuring' category and the remainder of the features is reassuring.
Pathological	A CTG whose features fall into two or more non-reassuring categories or one or more abnormal categories

Table 8 (above) helps to classify CTGs into one of the three groups where each group has a specific management plan.

If a CTG is normal, then the fetus is healthy and there is no specific action required.

In cases where the CTG is suspicious, intrapartum resuscitation should be instituted and CTG repeated.

If the CTG is pathological, intrapartum resuscitation and fetal blood sampling should be performed. In situations where fetal blood sampling is contraindicated or not possible then delivery should be expedited by an assisted vaginal delivery if the cervix is fully dilated or caesarian section if not imminently deliverable.[27]

When planning an intervention based on the changes on the CTG trace, the clinical picture as well as progress of labour needs to be taken into consideration.[2, 13, 28]

The use of a standard, universal classification of the CTG is a necessity .[23, 27, 29, 30] It helps the clinicians to understand and effectively communicate issues relating to fetal wellbeing in one common language. It also helps the clinicians to appropriately institute interventions to prevent perinatal morbidity and mortality.[29]

In South Africa there are no national or institutional guidelines for the interpretation of EFM traces. The result is that doctors do not know what to use especially in district hospitals where there are many inexperienced doctors.

1.8 Problem Statement

The implementation of guidelines is intended to improve the efficacy of EFM .The interpretation of the components of CTGs and NSTs vary from one clinician to the next, reducing the sensitivity of EFM.

So far, the NICE guidelines for NST/CTG interpretation have not been evaluated; neither have they been compared to non-systematic assessments by eyeballing. The implementation of a scoring system that visually tabulates the NICE guidelines may improve the efficacy of EFM interpretation and reduce inter- and intra-observer variability and attempt to attain uniformity of interpretation.

A tabular scoring system that uses the NICE guidelines to interpret CTGs has not been documented in the literature and hence the reason for this study.

2. OBJECTIVES

The objectives of this study are to:

1) Assess consensus in the interpretation of CTGs and NSTs between different grades of obstetric clinical staff by comparing assessment of traces by non-systematic eyeballing with assessment of traces using a tabular approach suggested by the NICE guidelines for interpretation of CTGs and NSTs.

2) Identify components of NSTs and CTGs with which obstetric personnel experience difficulty with interpretation.

3. METHODS

3.1 Setting

This study was conducted at the Charlotte Maxeke Johannesburg Academic, Rahima Moosa and Chris Hani Baragwanath Academic hospital maternity units. These are tertiary care hospitals for the teaching and training of the Witwatersrand university postgraduates and are major referral centers for the Johannesburg region. These institutions also contribute to the training of midwives and interns.

Registrars are continuously but informally trained on CTG interpretation through bedside teaching by consultants, tutorials, presentations and by attending perinatal morbidity and mortality meetings. Recently, a project by Johnson and Johnson in the Gauteng region has been dedicated to the formal training of midwives, interns and medical officers working in the maternity units on CTG interpretation.

Currently, the above mentioned maternity units make use of NSTs and CTGs for the evaluation of fetal wellbeing in high risk pregnancy and labour. NSTs and CTGs are performed in various wards by the nursing staff. The tracing is then presented to a doctor who interprets it and proposes further management of the patient. Fetal scalp blood sampling is not practiced due to the unknown rate of HIV seroconversion during pregnancy in the South African population, and because of practical difficulties with implementing fetal scalp blood sampling and analysis.

3.2 Study design and study population

This is a prospective two part questionnaire study, conducted from April 2010 to September 2010. The study population was midwives, advanced midwives, interns, medical officers, registrars and specialists working in Chris Hani Baragwanath Academic, Charlotte Maxeke Johannesburg Academic and Rahima Moosa hospitals, being the personnel who assess EFM traces. A convenience sample of 100 such staff members was recruited from these institutions, according to their availability and willingness to participate in the study.

3.3 Data Collection

Participants were recruited at the time of formal gatherings and departmental meetings in the various institutions on days when the researcher was available. They were informed about the study and signed informed consent to participate. To avoid contamination of data, participants were requested not to communicate with each other about the study.

3.3.1 The 2-part questionnaire

Each participant completed two separate questionnaires that were kept anonymous by allocating study numbers. An attendance register of all the participants was made to prevent duplication of participation.

3.3.2 First Questionnaire

The first questionnaire included the rank of the participant and a total of five traces to interpret and assess using the following three global assessment categories, based on the management plan.

1. Baby well
2. Baby needs further surveillance

3. Baby needs immediate delivery

The five traces were a combination of CTGs and NSTs. These traces were labeled 1-5. In order to maintain uniformity, the same five traces were presented to all participants as scanned images. No clinical information was presented with the traces.

First Questionnaire (see also appendix A)

Please mark appropriate rank with X

Rank	**X**
Midwife	
Advanced midwife	
Intern	
Community service	
Medical officer	
Registrar year 1 - 2	
Registrar year 3 - 4	
Senior medical officer	
Principal medical officer	
Specialist	

Please review each tracing and record your assessment according to one of the following assessments:

1. Baby well
2. Baby needs further surveillance
3. Baby needs immediate delivery

CTG/NST TRACE	**ASSESSMENT OF TRACE**
1	
2	
3	
4	
5	

3.3.3 Second Questionnaire

A second questionnaire was issued to each participant immediately after collection of the first questionnaire. The second questionnaire included the same traces as the first, but an adapted tabular model of the NICE guidelines was provided for the participants to score the traces systematically (Table 9). The five traces were presented in a different sequence from those in the first questionnaire. Participants were asked to score the five traces separately for baseline, baseline variability, decelerations and accelerations. As shown on Table 9, they were then asked to interpret each of the tracings as normal, suspicious or pathological.

Table 9 Tabular model of the NICE Guidelines with a scoring modification (see also appendix B)

Feature	**Mark appropriate box below with X**				**Number of parameters marked with X**	**Assessment of trace Ring appropriate number of features marked with X**	
	Baseline (BPM)	**Variability (BPM)**	**Decelerations**	**Accelerations**			
Reassuring	110-160	≥5	None	Present		**4X**	Normal
Non-reassuring	100-109 161-180	<5 for 40-90 minutes	Early decelerations. Variable decelerations with > 50% of contractions for > 90 minutes. Single prolonged deceleration for < 3 minutes.	No accelerations		**1X**	Suspicious
						≥2X	Pathological
Abnormal	<100 >180 Sinusoidal pattern > 10 minutes	<5 for > 90 minutes	Atypical variable decelerations with >50%of contractions for>30 minutes. Late decelerations for > 30 minutes. Single prolonged deceleration for > 3 minutes.			**≥1X**	Pathological

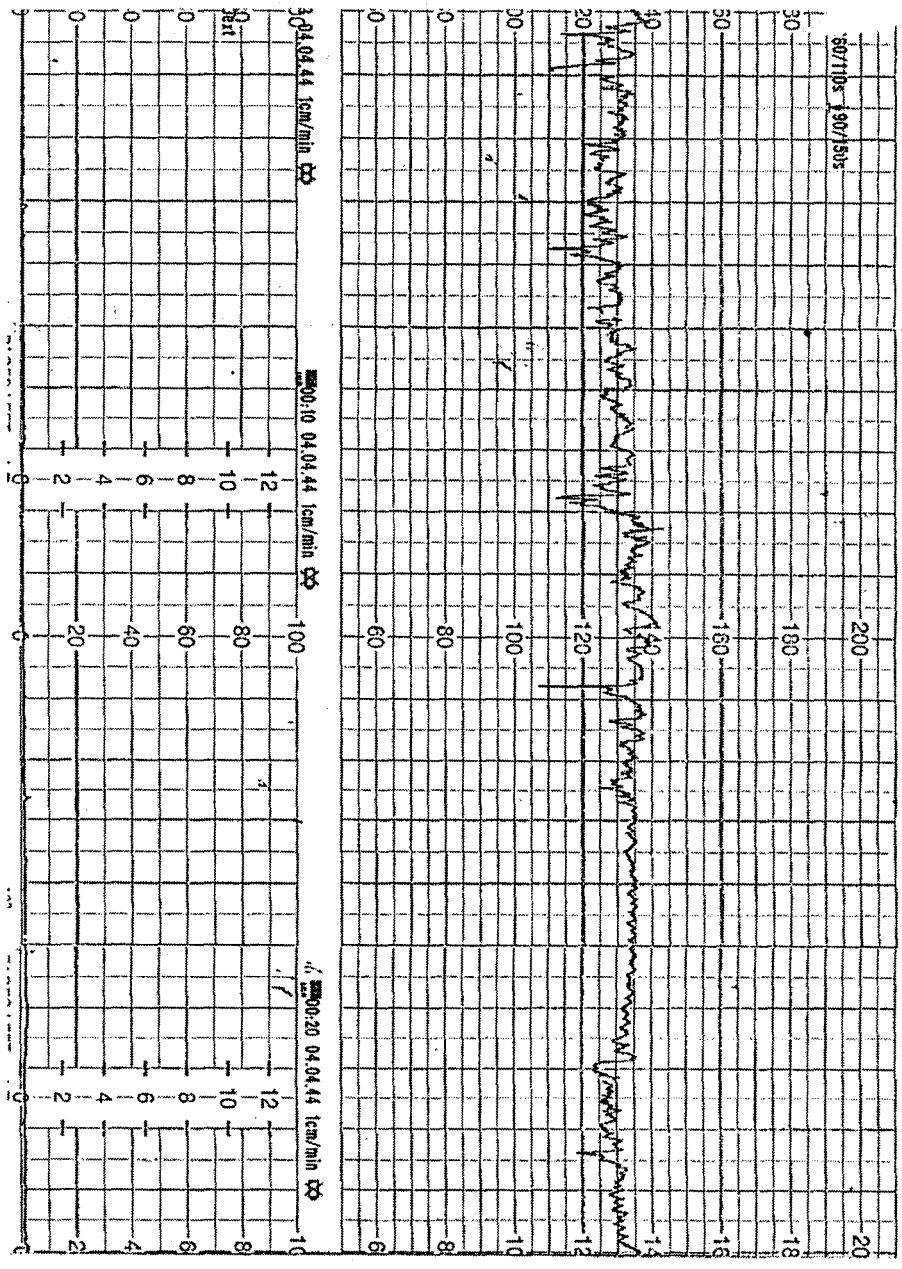


Figure 1 Trace 1

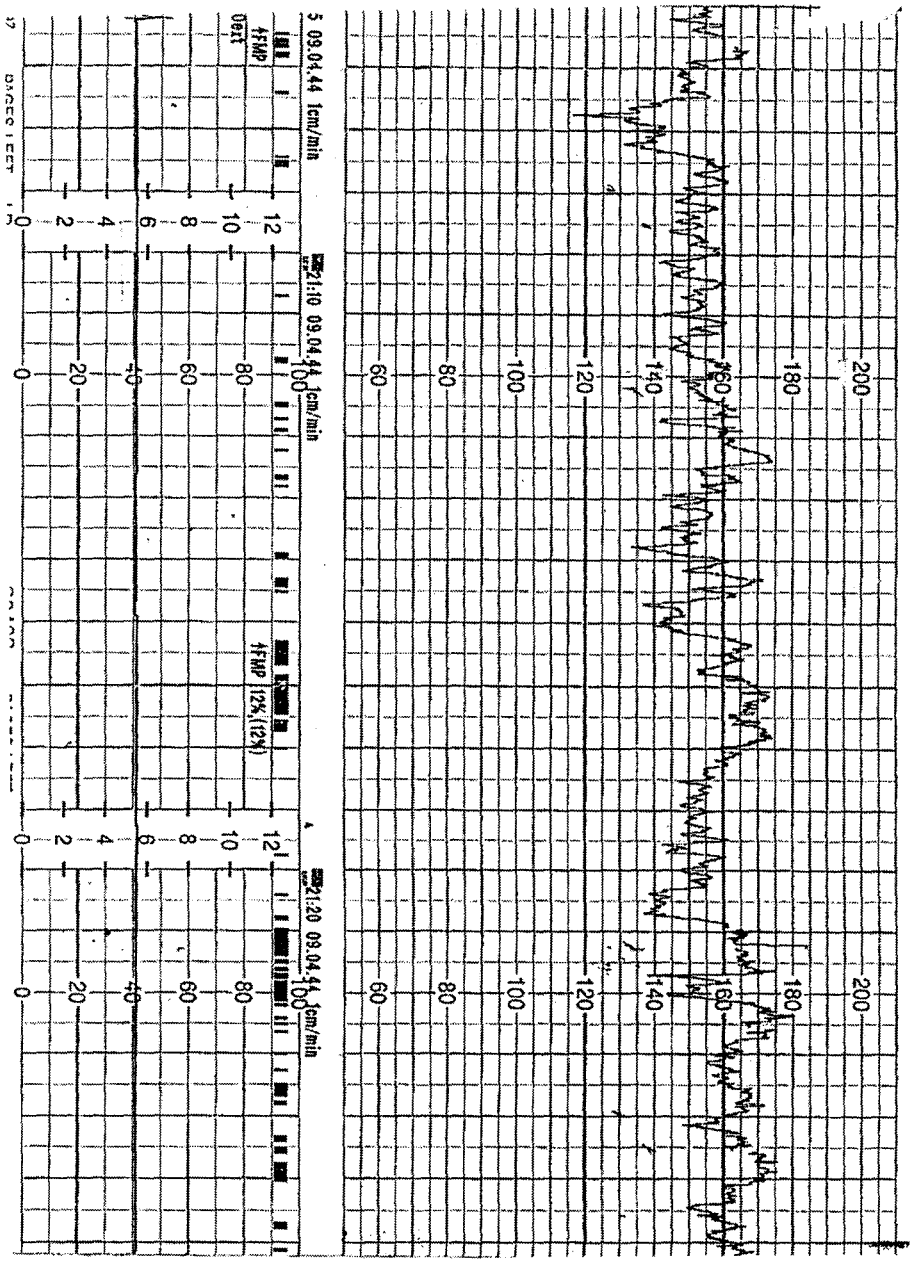


Figure 2 Trace 2

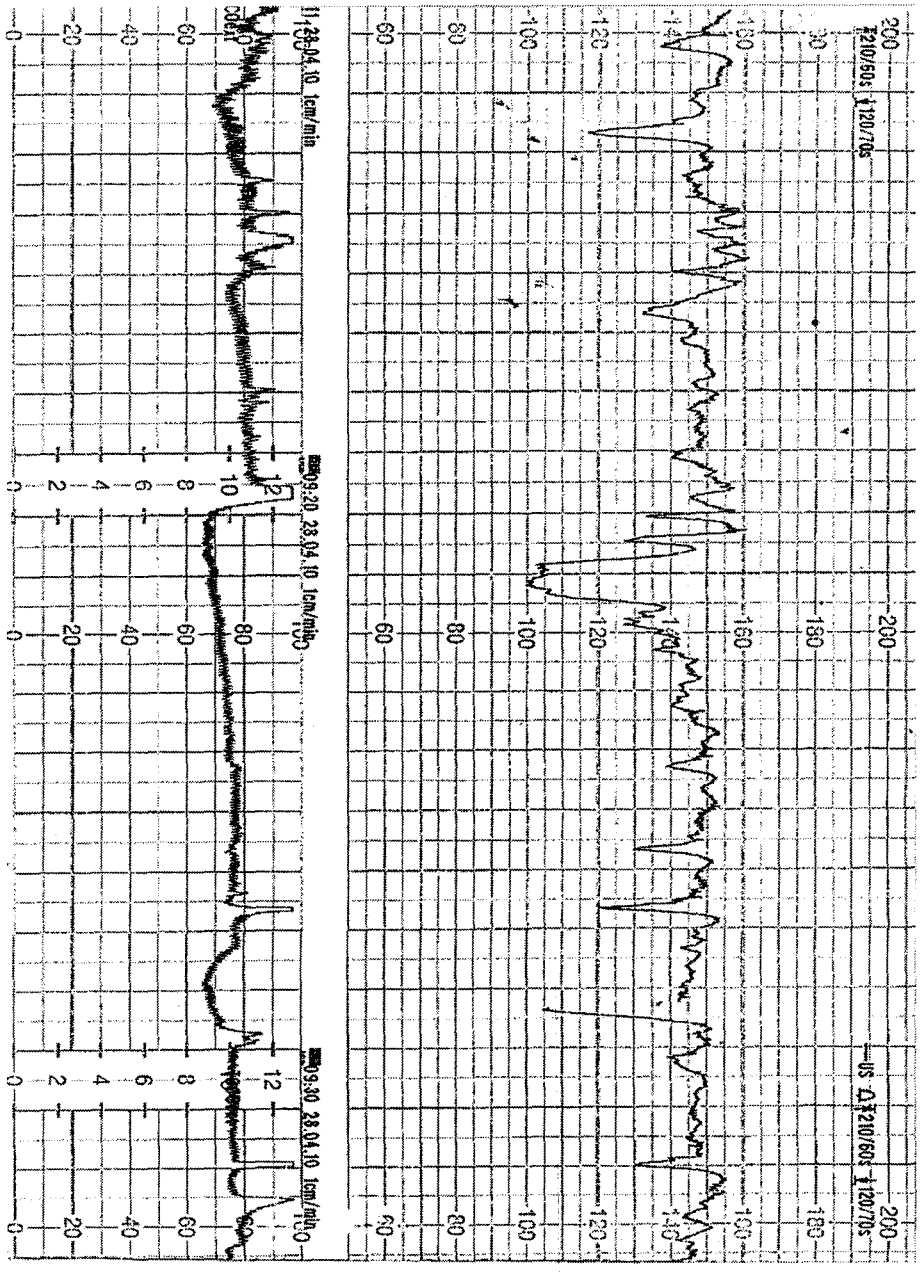


Figure 3 Trace 3

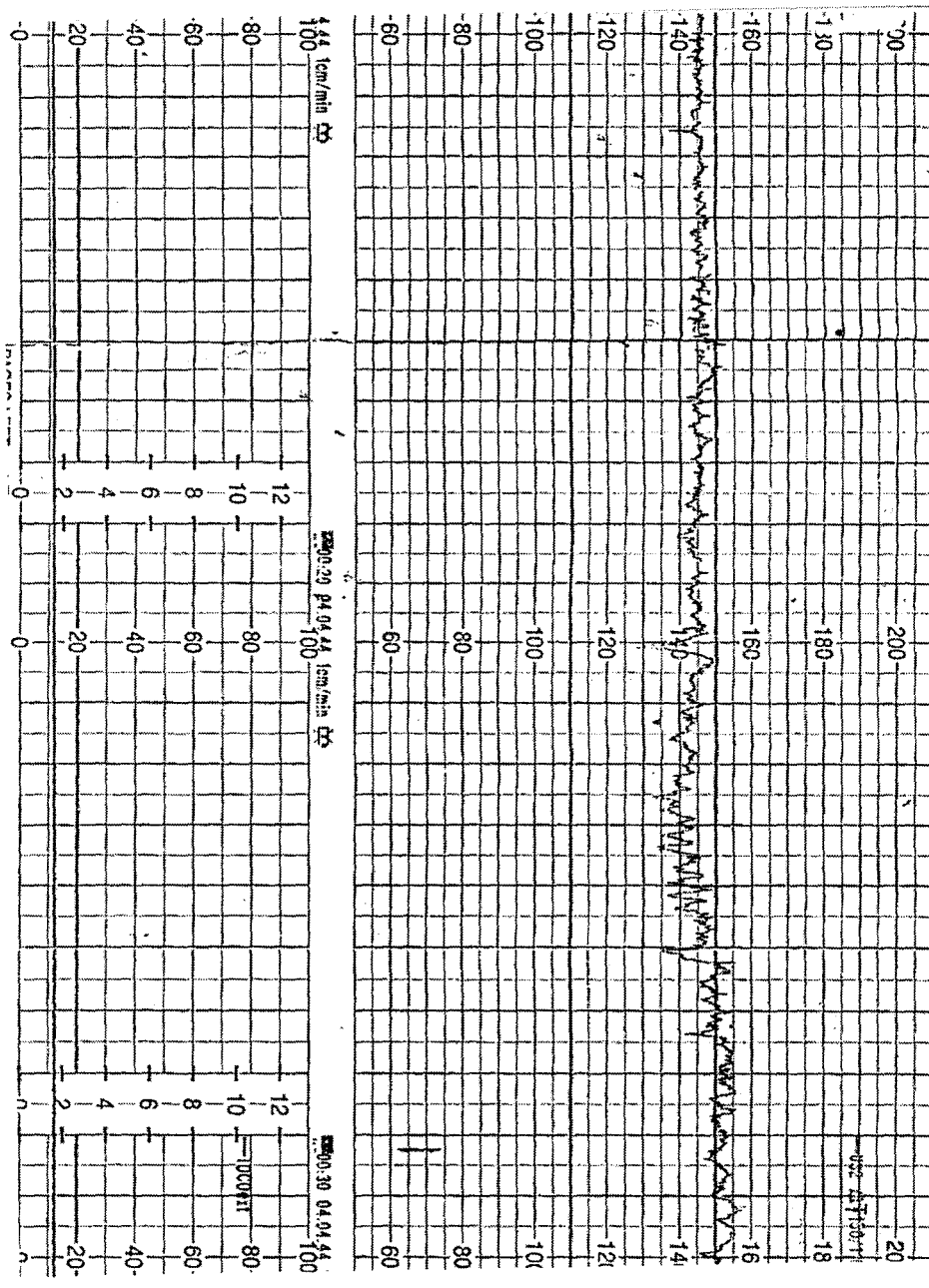


Figure 4 Trace 4

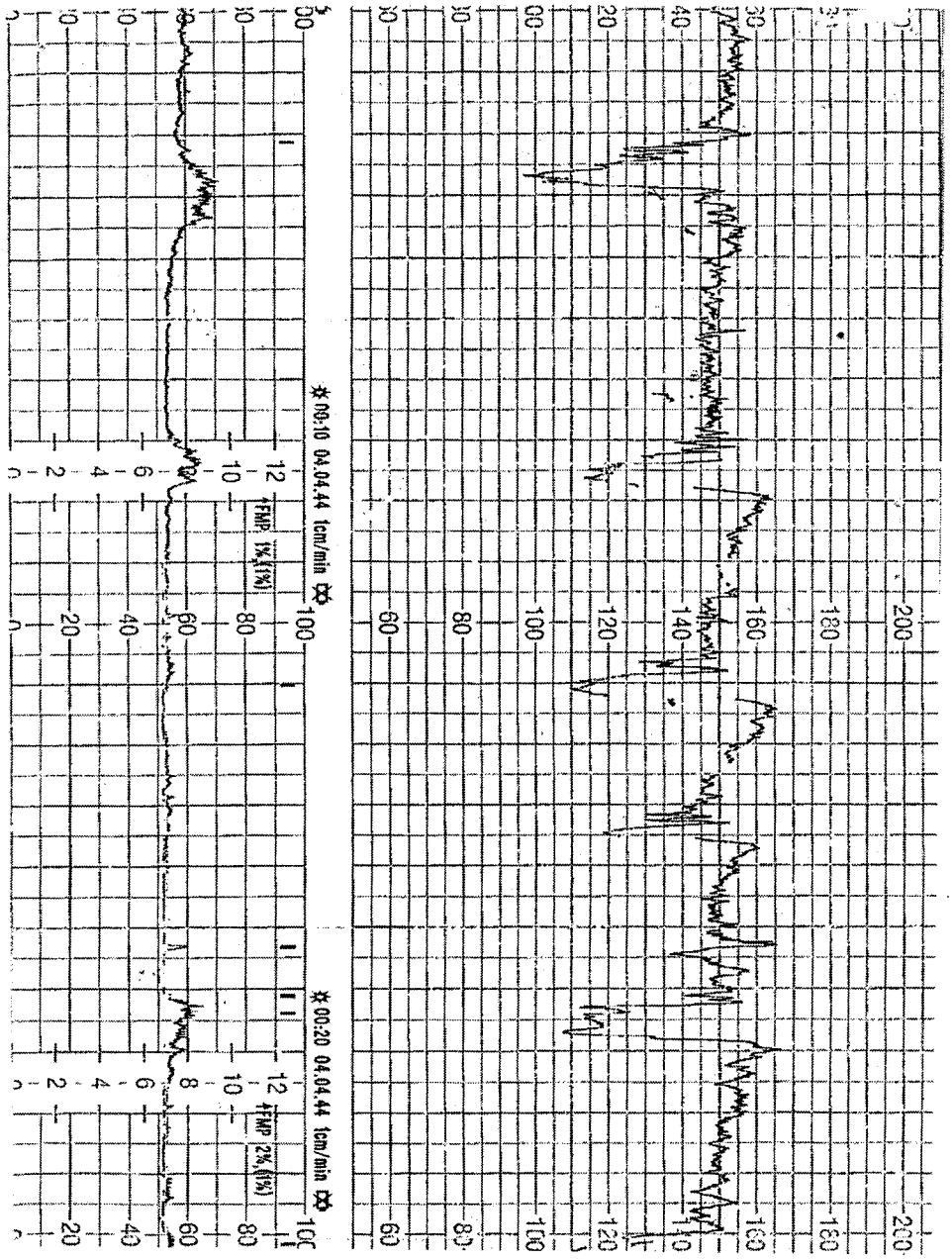


Figure 5 Trace 5

3.4 Data analysis

The study participants were categorized into three groups according to their rank where group1 constituted midwives, Group 2 - junior doctors (where junior doctors comprised interns, community service doctors, medical officers and registrars year 1 to year 2) and Group 3- senior doctors (registrars year 3 to year 4, senior medical officers, principal medical officers and specialists).

The data was managed and analyzed using Stata 11 software. Descriptive statistics were employed by using means with standard deviations and medians with ranges.

Frequencies were expressed in percentages with 95% confidence intervals. Differences in interpretation of each of the tracings by rank of health personnel (midwives, junior doctors, senior doctors) were assessed using Fisher's exact test. Statistical significance was assumed at a P-value less than 0.05. In order to determine the degree of certainty between study participants, a 75% consensus among participants was used to determine certainty of the evaluation. If 75% or more of the participants agreed on a category of evaluation then the evaluation by the study participants was considered to be certain. If consensus of participants was below 75% then there was uncertainty about the evaluation.

3.5 Ethics

The study was commenced after obtaining an ethical approval from the University of Witwatersrand Ethics Committee, M091136.

Consent was obtained from each participant prior to inclusion in the study. In order to maintain anonymity, each participant was allocated a study number. The data sheet only reflected data

pertaining to the study and did not include the date, name or hospital of the participant.

Participants retained the right to withdraw from the study at any time.

The NSTs and CTGs did not reflect the patients' names or hospital numbers.

4. RESULTS

There was a total of 100 participants who were medical staff ranging from midwives to specialist obstetricians and gynecologists. Table 10 reflects the rank of the study participants.

Table 10 Medical rank of participants and grouping into midwives, junior doctors and senior doctors

RANK	n	RANKGROUP	n
Midwife	12	Midwives (Group 1)	15
Advanced midwife	3		
Intern	10	Junior doctors (Group 2)	33
Community service	4		
Medical officer	4		
Registrar year 1-2	15		
Registrar year 3-4	10	Senior doctors (Group 3)	50
Senior medical officer	4		
Principal medical officer	9		
Specialist	27		
Blank	1	Excluded	2
Incorrect	1		
Total	100		

The study participants were categorized into three groups as described earlier. There were 15 study participants in group1, 33 in group 2 and 50 in group 3 (Table 10).

The non-systematic assessment of the five traces by the participants not grouped by grade of staff is represented in Table 11. The nonsystematic evaluation was correctly completed by 98%

of participants for Traces 1 and 2. Ninety-seven percent of study participants correctly completed Traces 3, 4 and 5. The non-systematic assessment yielded varying interpretations by the various categories of medical staff. Interpretation of Trace 3 gave the best agreement, with 77% of personnel stating that the baby needed immediate delivery.

Table 11 Non systematic assessment of traces (n = 100)

	TRACE 1	TRACE 2	TRACE 3	TRACE 4	TRACE 5
Baby well	2	30	1	38	32
Baby requires further surveillance	28	66	19	53	63
Baby needs immediate delivery	68	2	77	6	2
Blank	1	1	1	1	1
Incorrectly completed	1	1	2	2	2
Total	100	100	100	100	100

Table 12 displays the interpretation of traces by the same participants using the tabular approach adapted from the NICE guidelines on electronic fetal monitoring.

Good agreement on the description of the baseline was achieved in all five traces. Good agreement on the classification of baseline variability was seen in Traces 1, 4 and 5. The interpretation of baseline variability in Traces 2 and 3 varied, with the majority of participants describing the baseline variability as reassuring. The number of study participants who did not complete the baseline assessment ranged from 6% to 9% of the various traces.

The interpretation of decelerations varied among the participants. There was good agreement in Traces 1 and 2 where 78% and 90% respectively agreed on the interpretation of decelerations. There was good agreement on the recognition of accelerations in Traces 2, 4 and 5.

The overall assessment of the traces varied among the participants as to whether the traces were

normal, suspicious or pathological. Only in Trace 3 was there good agreement (84%). A large number of study participants did not complete the overall assessment of the trace, ranging from 21% to 24% of the various traces, and up to 5% were incorrectly completed.

Table 12 Systemic evaluation of traces using the NICE guidelines (n = 100)

		TRACE 1	TRACE 2	TRACE 3	TRACE 4	TRACE 5
BASELINE						
Reassuring		89(98)	93(100)	91(100)	82(90)	92(100)
Non-reassuring		2(2)	0	0	7(8)	0
Abnormal		0	0	0	2(2)	0
Blank		9	6	8	8	7
Incorrect		0	1	1	1	1
BASELINE VARIABILITY						
Reassuring		78(87)	60(66)	46(52)	84(92)	76(84)
Non-reassuring		10(11)	29(32)	37(42)	6(7)	14(15)
Abnormal		2(2)	2(2)	5(6)	1(1)	1(1)
Blank		10	9	12	9	9
Incorrect		0	0	0	0	0
DECELERATIONS						
Reassuring		0	81(90)	0	36(42)	56(64)
Non-reassuring		70(78)	9(10)	42(49)	42(48)	28(32)
Abnormal	Atypical variable	13(14)	0	18(21)	7(8)	3(3)
	Late	4(4)	0	19(22)	2(2)	0
	Single prolonged >3min	3(3)	0	7(8)	0	0
Blank		7	10	6	10	12
Incorrect		3	0	8	3	1
ACCELERATIONS						
Reassuring(present)		54(63)	7(8)	31(35)	89(96)	12(14)
Non-reassuring(absent)		31(37)	85(92)	57(65)	4(4)	76(86)
Blank		12	8	9	7	10
Incorrect		3	0	3	0	2
ASSESSMENT						
Normal		0	7	0	36(46)	11(15)
Suspicious		37(47)	49(69)	12(16)	36(46)	46(62)
Pathological		41(53)	22(31)	62(84)	7(8)	17(23)
Blank		19	21	21	21	24
Incorrect		3	1	5	0	2

*Calculated percentages in brackets excluding those that were unknown.

Table 13 compares the non-systematic assessment of traces according to the grades of staff. There were statistically significant differences in the assessments of traces between the groups of participants in Traces 1, 2 and 4.

In Trace 1, 28 of the junior doctors (85%) and 11 of the midwives (73%) suggested immediate delivery, while only 27 senior doctors (56%) made a similar suggestion (P = 0.005).

In Trace 2 (Table 13) senior doctors showed certainty in suggesting further surveillance, compared with a much lower proportion of midwives and junior doctors (P = 0.033).

There was a difference in the assessment of Trace 4 by senior doctors where the majority made a decision of "baby well", while the majority of junior doctors and midwives suggested further surveillance (P = 0.003).

The systematic evaluation of Traces 1 to 5 (Table 14), by implementation of the tabular approach to the NICE guidelines, revealed no statistically significant differences in the interpretation of the baseline and the assessment of the traces.

There was no statistically significant difference in the evaluation of each component (Baseline, Baseline variability, Decelerations, Accelerations, and Assessment) in Traces 2, 4 and 5 (Table 14).

A statistically significant difference in the evaluation of baseline variability was noted in Traces 1 and 3 where in Trace 1 the midwife group was uncertain (Table 14). In Trace 3 there was a statistically significant difference noted in the evaluation that was made by study participants in the groups and there was uncertainty within each group (Table 14).

There was a statistically significant difference in the evaluation of accelerations in Trace 1 where the midwife group was certain about the evaluation and both groups of doctors were uncertain (Table 14).

The evaluation of decelerations (Table 14) showed a statistically significant difference in Trace 3 where certainty was not achieved by all groups with respect to the type of deceleration. Fifty three percent of midwives evaluated the decelerations as "late" whereas the majority of junior and senior doctors evaluated the decelerations as "non-reassuring".

Table13 Non-systematic assessment of traces according to the grades of staff (n = 96)

C = CERTAIN (≥ 75% agreement by participants on trace interpretation, shown as an asterisk*.)

UC = UNCERTAIN (< 75% agreement by participants on trace interpretation.)

	TRACE1				TRACE2				TRACE3				TRACE4				TRACE5			
	Midwives (n =15)	Junior Doctors (n =33)	Senior doctors (n = 48)	P-VALUE	Midwives (n = 15)	Junior Doctors (n = 3)	Senior doctors (n=48)	P-VALUE	Midwives (n = 14)	Junior Doctors (n = 33)	Senior doctors (n = 48)	P-VALUE	Midwives (n = 14)	Junior Doctors (n = 33)	Senior doctors (n = 48)	P-VALUE	Midwives (n = 14)	Junior doctors (n = 33)	Senior doctors (n = 48)	P-VALUE
Baby Well	1	1	0	0.005	7	13	9	0.033	0	1	0	0.145	3	7	28	0.003	8	8	14	0.204
Baby requires further surveillance	3	4	21		7	20	38(79)*		2	3	13		10	23	19		6	24	33	
Baby needs immediate delivery	11	28(85)*	27		1	0	1		12(86)*	29(88)*	35		1	3	1		0	1	1	
Degree of certainty	UC	C	UC		UC	UC	C		C	C	UC		UC	UC	UC		UC	UC	UC	

Table 14 Systematic evaluation using NICE guidelines according to the grades of staff (n = 98)

C = CERTAIN (≥ 75% agreement by participants on trace interpretation, shown as an asterisk*.)

UC = UNCERTAIN (< 75% agreement by participants on trace interpretation)

	TRACE 1				TRACE 2				TRACE 3				TRACE 4				TRACE 5			
	Midwives n=15	Junior doctors n=33	Senior doctors n=50	P-value	Midwives n=15	Junior doctors n=33	Senior doctors n=50	P-value	Midwives n=15	Junior doctors n=33	Senior doctors n=50	P-value	Midwives n=15	Junior doctors n=33	Senior doctors n=50	P-value	Midwives n=15	Junior doctors n=33	Senior doctors n=50	P-value
BASELINE																				
Reassuring	14*	27*	47*		14*	29*	48*		14*	28*	48*		10	25*	46*		14*	28*	49*	
Non-reassuring	0	0	1	NS	0	0	0	NS	0	0	0	NS	3	2	2	NS	0	0	0	NS
Abnormal	0	0	0		0	0	0		0	0	0		1	0	1		0	0	0	
Degree of certainty	C	C	C		C	C	C		C	C	C		UC	C	C		C	C	C	
BASELINE VARIABILITY																				
Reassuring	7	23*	47*		9	16	33		3	20	21		8	25	48		11*	24*	40*	
Non-reassuring	5	3	1	<0.001	4	10	14	NS	6	7	23	0.006	3	2	1	NS	1	4	7	NS
Abnormal	1	0	0		0	1	0		3	1	1		1	0	0		1	0	0	
Degree of certainty	UC	C	C		UC	UC	UC		UC	UC	UC		UC	UC	UC		C	C	C	
DECELERATIONS																				
Reassuring	0	0	0		12*	25*	43*		0	0	0		5	9	22		4	19	33	
Non-reassuring	11*	17	41*		2	1	5		4	11	26		6	13	22		8	7	13	
Abnormal Atypical variable	1	7	5		0	0	0		1	6	10		1	3	3		1	1	1	
Abnormal-Late	1	0	3	NS	0	0	0	NS	8	6	5	0.037	0	2	0	NS	0	0	0	NS
Single prolonged >3minutes	1	2	0		0	0	0		1	3	3		0	0	0		0	0	0	
Degree of certainty	C	UC	C		C	C	C		UC	UC	UC		UC	UC	UC		UC	UC	UC	
ACCELERATIONS																				
Reassuring(present)	12*	18	24		3	0	4		3	12	16		15*	26*	47*		3	4	5	
Non-reassuring(absent)	1	8	22	0.02	12*	27*	44*	NS	12*	15	28	NS	0	2	2	NS	11*	22*	41*	NS
Degree of certainty	C	UC	UC		C	C	C		C	UC	UC		C	C	C		C	C	C	
ASSESSMENT																				
Normal	0	0	0		2	0	5		0	0	0		5	9	22		2	5	4	
Suspicious	6	10	21	NS	6	14	29	NS	1	3	8	NS	4	11	20	NS	5	11	30	NS
Pathological	7	11	22		4	8	10		11	18	33		3	3	1		3	6	8	
Degree of certainty	UC	UC	UC		UC	UC	UC		UC	UC	UC		UC	UC	UC		UC	UC	UC	

5. DISCUSSION

The primary objective of this study was to assess the uniformity of the observers at assessing NSTs and CTGs using the tabular approach from the NICE guidelines. Uniformity in the interpretation of NSTs and CTGs may aid in the prediction of fetal outcomes.

5.1 Medical rank of participants

The study sample in Table 10 displays a good representation of the clinical practitioners in the maternity units including midwives, junior doctors and senior doctors. However, this was a convenience sample, and the staff who agreed to participate may not be representative of all staff in the obstetric departments. In the literature, there is no documentation on the comparison of the different grades of staff in an obstetric unit with respect to interpretation of NSTs and CTGs.

It is a known fact that formal education and training in EFM improves CTG interpretation. Randomized controlled trials have suggested that clinical experience and training in EFM and interpretation of CTG traces improves knowledge and clinical skills.[28, 29]It is, however, not known whether this has an impact on maternal and neonatal outcomes. In 2005 Blix et al. demonstrated that there was good agreement in the interpretation of 549 labour admission tests by midwives and obstetricians who had completed a training program in "fetal surveillance".[29]

5.2 The non-systematic assessment of traces

Table 11 shows considerable uncertainty in the non-systematic assessment of the traces with a statistically significant difference in the assessment of 3 of the traces between the groups as evidenced in Table 13.

There is no uniformity in this assessment of traces and this is demonstrated by the fact that the degree of certainty varied among the staff grades. The non-systematic assessment of traces may incorporate a guideline approach for some individuals but also includes a component of experience which may have influenced the assessment.

5.3 Systematic evaluation of traces using the NICE guidelines tabular model with a scoring modification

5.3.1 Interpretation of fetal heart rate components

The number of study participants who evaluated the traces by assessing the various components of a trace was reduced by the use of the tabular approach from the NICE guidelines. Table 11 shows that 1% of the study participants did not complete the non-systemic assessment. When the findings of Table 11 are compared with Table 12 where 6% - 12% (Table 12) of study participants did not complete fetal heart rate components and 19% - 24% (Table 12) of the overall assessments were not completed. This may indicate that the participants were not familiar with the NICE guidelines approach or that they did not know how to apply information contained within the guidelines. It was also observed that some of the study participants evaluated each fetal heart rate component but were not able to assimilate the information into an overall assessment based on the contribution of the four features.

The findings revealed in Table 12 show that there was general agreement among participants on the classification of the baseline FHR. The study participants demonstrated uniformity in the assessment of baseline, which means that they found it uncomplicated to interpret. This finding is in keeping with the previous studies.[2, 8, 17, 18, 19] Ayres-de-Campos et al. conducted a study in 2004 which demonstrated that the interpretation of fetal heart rate baseline is extremely reproducible.[28] A striking phenomenon in his study was the fact that agreement among the clinicians was higher if the clinicians were fully familiarized with the criteria used for classification and had prior training.

The interpretation of baseline variability, decelerations and accelerations varied among the participants. The variation and uncertainty in the interpretation of baseline variability is in keeping with the previous studies as baseline variability was found to be the most difficult to interpret. The diversity in the evaluation of baseline variability, decelerations, the type of deceleration and accelerations demonstrates difficulty in the interpretation of these components. There has been inconsistency with the findings on the assessment of decelerations and accelerations in the literature. Difficulty in the evaluation of accelerations, decelerations and baseline variability in this study could be explained by: vague definitions in the assessment of fetal heart tracings; the nature of the FHR pattern (the more abnormal the FHR pattern the more difficult the interpretation); lack of experience and training on CTG analysis; short CTG tracings. This was, however, not measured in the study.

Certainty in the recognition of accelerations was much better among the study participants when compared with the interpretation of decelerations and baseline variability. This may mean that

the study participants understand the definition of accelerations better than the other two CTG components.

5.3.2 Classification of fetal heart rate patterns using the NICE guidelines

The overall assessment of traces, using the NICE guidelines varied between the study participants. There is a statistically significant difference displayed among the rank groups in the non-systematic assessment of the traces as seen in Traces 1, 2 and 3 (Table 13). There is, however, no statistically significant difference noted among the groups in all the traces (Table 14), when the NICE guidelines tabular approach is used to interpret NSTs and CTGs. Significant differences in interpretation of CTGs by the different grades of staff on non-systematic assessment disappeared when the tabular NICE definitions were applied.

This means that the different rank groups interpret the traces more uniformly when using the NICE guidelines tabular approach, suggesting that staff grade-related differences in trace interpretations are reduced. However, there was still uncertainty in all the groups even with the NICE guidelines approach.

5.3.3 Problem areas identified in the interpretation of NSTs and CTGs

The problem areas highlighted by this study are in the interpretation of the baseline variability, decelerations and the recognition of accelerations. A simpler system to define and interpret these fetal heart rate components, and a consensus regarding the management of specific patterns are needed. More work, by the maternity units, needs to be done to educate and train medical staff on CTG interpretation. The use of computerized CTG analysis may also assist with consistent

assessment of fetal heart rate patterns, but such modalities have not yet been fully evaluated in trials and meta-analysis.

5.3.4 Limitations of this study

1. The NST and CTG traces used in the study were short (20 minute traces). This makes the classification difficult especially for the interpretation of baseline variability which may need up to a 90-minute trace.
2. Clinical information was not given to the observers with the NSTs and CTGs and this may add insight to the interpretation of a trace. The objectives of this study were not about the accuracy of CTG interpretation, it was about the consensus in the interpretation of CTGs among the different grades of obstetrics staff.
3. The study did not evaluate the participant's prior education or level of training on CTG interpretation and it is known that the assessment of uniformity in CTG interpretation is influenced by experience or the level of training.[28, 29] Senior staff members are well trained on CTG interpretation but may not be familiar with the NICE guidelines approach as there is no standard protocol on which system to use for interpreting CTGs by these different institutions.
4. The sample size may be too small. It may have shown no significant difference where there was in fact a difference (type 2 error), especially in the comparison of the NICE tabular approach.

6. CONCLUSION

Some uniformity in the interpretation of CTGs and NSTs and reduction in inter-observer variation is attained by the use of the NICE guidelines tabular approach.

There is a need for a standard, universal classification of CTGs. This has been addressed by many researchers as an intervention to try and minimize inter-observer variation in the interpretation of CTGs. The NICE guidelines tabular approach provides a system with a standardized terminology; a system that is simple to use, applicable to clinical practice, and reproducible. Education and training of the clinicians in maternity units on CTG interpretation is crucial. This would probably best be performed by individuals well versed in CTG interpretation as this study has also shown that there is a wide variation in interpretation which is also influenced by the clinical experience of the clinician.

The application of the NICE guidelines may be better facilitated by using the tabular approach in the form of a stamp on each CTG or NST, where the clinician needs to classify each of the fetal heart rate features and then gives the overall assessment of the tracing.

7. REFERENCES

1. Royal College of Obstetricians and Gynaecologists. The use of electronic fetal monitoring. The use and interpretation of CTG in intrapartum fetal surveillance. Evidence-based Clinical Guideline No. 8. London: RCOG, 2001.

2. Donker DK, VanGeijn H, Hasman A. Interobserver variation in the assessment of fetal heart rate recordings. Eur J Obstet Gynaecol Reprod Biol 1993;52:21-28.

3. Gibb D, Arulkumaran S. Fetal monitoring in practice, 2nd edition. Oxford: Butterworth-Heinemann. 1997;44-119.

4. Impey L, Reynolds M, MacQuillan K, Gates S, Murphy J, Sheil O. Admission cardiotocography: a randomized controlled trial. The Lancet 2003;361:465-470.

5. Devane D, Lalor J. Midwives' visual interpretation of intrapartum cardiotocographs: intra- and inter-observer agreement. J Adv Nurs 2005;52:133-141.

6. Graham E, Petersen S, Christo D, Fox H. Intrapartum electronic fetal heart rate monitoring and the prevention of perinatal brain injury. Obstet Gynecol 2006;108:656-666.

7. Schiermeier S, Pildner van Steinburg S, Thiema A, Reinhard J, Daumer M, Scholz M, Hatzmann W, Schneider K. Sensitivity and specificity of intrapartum computerized FIGO criteria for cardiotocography and fetal scalp pH during labour: multicentre, observational study. BJOG 2008;115:1557-1563.

8. Alfirevic Z, Devane D, Gyte GM. Continuous tocography(CTG) as a form of electronic fetal monitoring(EFM) for fetal assessment during labour. Cochrane Database Syst Rev 2006; 3:CD006066.

9. Premila S, Arulkumaran S. Intrapartum fetal surveillance. Obstet Gynaecol Reprod Med 2007;18:12-17.

10. Amer-Wahlin I, Yli B, Arulkumaran S. Foetal ECG and STAN technology – a review. Eur Clin Obstet Gynaecol 2005;1:61-73.

11. Williams B, Arulkumaran S. Cardiotocography and medicolegal issues. Best Pract Res Clin Obstet Gynaecol 2004;18:452-466.

12. Bernardes J, Costa-Pereira A, Ayres-de-Campos D, van Geijn H. P, Pereira-Leite L. Evaluation of interobserver agreement of cardiotocograms. Int J Gynecol Obstet 1997;57:33 - 37.

13. Trimbos J.B, Keirse M.J.N.C. Observer variability in assessment of antepartum cardiotocograms. BJOG 1978;85:900-906.

14. Mires G, Williams F, Howie P. Randomised controlled trial of cardiotocography versus doppler auscultation of fetal heart at admission in labour in low risk obstetric population. BMJ 2001;322:1457-1460.

15. Lotgering F.K, Wallenburg H, Schouten H.J. Interobserver and intraobserver variation in the assessment of antepartum cardiotocograms. Am J Obstet Gynecol 1982;144:701-705.

16. Todros T, Prere CU, Plazzolta C, Biolcati M, Lombardo P. Fetal heart rate tracings: observers versus computer assessment. Eur J Obstet Gynecol Reprod Biol 1996;68:83-86.

17. National Institute of Child Health and Human Development Research Planning Workshop. Electronic fetal heart rate monitoring: Research guidelines for interpretation. Am J Obstet Gynecol 1997;177:1385-1390.

18. Blix E, Sviggam O, Koss KS, Qian P. Interobserver variation in assessment of 845 labour admission tests:comparison between midwives and obstetricians in the clinical setting and two experts. BJOG 2003;110:1-5.

19. ACOG Practice Bulletin No.106. Intrapartum fetal heart rate monitoring: Nomenclature, interpretation and general management principles. Obstet Gynecol 2009;114:192-202.

20. Ayres-de-Campos D, Bernardes J, Costa-Pereira A, Pereire-Leite L. Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. BJOG 1999;106:1307-1.

21. Ayres-de-Campos D, Bernardes J. Twenty five years after the FIGO guidelines for the use of fetal monitoring: Time for a simplified approach? Int J Gynecol Obstet 2010;110:1-6.

22. FIGO subcommittee on standards in Perinatal medicine. Guidelines for the use of fetal monitoring. Int J Gynecol Obstet 1987;25:159-167.

23. Robinson B, Nelson L. Update on definitions, interpretative systems with management strategies, and research priorities in relation to intrapartum electronic fetal monitoring. Rev Obstet Gynecol 2008;1:186-192.

24. Chandraharan E, Arulkumaran S. Prevention of birth asphyxia: responding appropriately to cardiotocograph traces. Best Pract Res Clin Obstet Gynaecol 2007;21:609-624.

25. Arulkumaran S, Symonds EM. Intrapartum fetal monitoring-medico-legal implications. The Obstetrician and Gynaecologist 1999;1:23-26.

26. Robinson B. A review of NICHD standardized nomenclature for cardiotocography: The importance of speaking a common language when describing electronic fetal monitoring. Rev Obstet Gynecol 2008;1:56-60.

27. Liston R, Crane J, Hughes O, Kuling S, MacKinnon C, Milne K, Richardson B, Trepanier MJ, et al. Fetal health surveillance in labour. J Obstet Gynaecol Can 2002;24:250-276.

28. Ayres-de-Campos D, Bernardes J, Marsal K, Nickelsen C, Makarainen L, Banfield P, Xavier P, Campos I. Can the reproducibility of fetal heart rate baseline estimation be improved? Eur J Obstet Gynaecol Reprod Biol 2004;112:49-54.

29. Blix E, Qian P. Interobserver agreements in assessing 549 labor admission tests after a standardized training program. Acta Obstet Gynecol Scand 2005;84:1087-1092.

30. Parer JT, King T. Fetal heart rate monitoring: is it salvageable? Am J Obstet Gynecol 2000;182:982-987.

APPENDIX A First questionnaire

Please mark appropriate rank with X

Rank	X
Midwife	
Advanced midwife	
Intern	
Community service	
Medical officer	
Registrar year 1 - 2	
Registrar year 3 - 4	
Senior medical officer	
Principal medical officer	
Specialist	

Please review each tracing and record your assessment according to one of the following assessments:

1. Baby well
2. Baby needs further surveillance
3. Baby needs immediate delivery

CTG/NST TRACE	ASSESSMENT OF TRACE
1	
2	
3	
4	
5	

APPENDIX B Second questionnaire

Feature	Mark appropriate box below with X				Number of parameters marked with X	Assessment of trace Ring appropriate number of Features marked with X	
	Baseline (BPM)	Variability (BPM)	Decelerations	Accelerations			
Reassuring	110-160	≥ 5	None	Present		**4X**	Normal
Non-reassuring	100-109 161-180	< 5 for ≥ 40mins but< 90 minutes	Early decelerations Variable decelerations with > 50% of contractions for > 90mins Single prolonged deceleration for < 3mins.	No accelerations		**1X**	Suspicious
						≥2X	Pathological
Abnormal	<100 >180 Sinusoidal pattern ≥ 10mins.	<5 for ≥ 90 minutes	Atypical variable decelerations with >50% of contractions, > 30 minutes Late decelerations, > 30 minutes Single prolonged decelerations for > 3 minutes			**≥1X**	Pathological

APPENDIX C Ethics approval

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Dr N.B. Jack

CLEARANCE CERTIFICATE	**M091136**
PROJECT	The Evaluation of a Tabular Application of the National Institute of Health and Clinical Excellence (NICE) Guidelines for Universal Interpretation of Non-Stress Tests (NST) and Cardiotocographs (CTG)
INVESTIGATORS	Dr N.B. Jack.
DEPARTMENT	Obstetrics & Gynaecology
DATE CONSIDERED	2009/11/27
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 2009/11/30 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable

cc: Supervisor : Dr P Naidoo

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...