From Department of Women's and Children's Health
Karolinska Institutet, Stockholm, Sweden

DIFFERENT ASPECTS OF ELECTRONIC FETAL MONITORING DURING LABOR

Erika Gyllencreutz

Karolinska Institutet

Stockholm 2022

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-807-6
Cover illustration: Elisabeth Abdon

Different aspects of electronic fetal monitoring during labor

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Erika Gyllencreutz

The thesis will be defended in public at Skandiasalen, QA31:01, Karolinska University Hospital, Solna, Karolinska vägen 37A, Stockholm, Friday 2nd of December 2022 at 09.00 am

Principal Supervisor:
Docent Malin Holzmann
Karolinska Institutet
Department of Women's and Children's Health
Division of Neonatology, Obstetrics and Gynecology

Co-supervisor(s):
Professor Pelle Lindqvist
Karolinska Institutet
Department of Clinical Science and Education
Division of Obstetrics and Gynecology

PhD Ingela Hulthén Varli
Karolinska Institutet
Department of Women's and Children's Health
Division of Neonatology, Obstetrics and Gynecology

Opponent:
Professor Marie Blomberg
Linköping University
Department of Biomedical and Clinical Sciences
Division of Children's and Women's Health

Examination Board:
Docent Ulf Hanson
Uppsala University
Department of Women's and Children's Health,
Division of Obstetric and Reproductive Health Research

Docent Anna Sandström
Karolinska Institutet
Department of Medicine
Division of Clinical Epidemiology

Docent Kajsa Bohlin
Karolinska Institutet
Department of Clinical Science, Intervention and Technology
Division of Paediatrics

To the beloved small and big men in my life:

Viktor, August, Vilhelm and Christian

POPULAR SCIENCE SUMMARY OF THE THESIS

Cardiotocography, or CTG for short, was developed in the 1960's as a means to assess the wellbeing of the fetus during labor and subsequently introduce the possibility to intervene during labor when signs of fetal distress (shortage of oxygen) arose. The CTG machine consists of devices that continuously measure the fetal pulse rate and the contractions of the uterus concomitantly. Common interventional tools, when signs of fetal distress occur, are cesarean and instrumental delivery, most commonly performed by vacuum extraction. Nonetheless, the interventional procedures are operative and as such associated with sometimes severe side effects. CTG on the other hand has been criticized as a blunt diagnostic instrument – abnormal CTG tracings are indeed very common during labor even when the fetus is perfectly well oxygenated. Consequently, the use of CTG has been associated with increased risk of cesarean deliveries without proven effect on perinatal mortality or cerebral palsy. Nonetheless, to detect differences in these severe and very rare outcomes, extremely large studies are necessary and cerebral palsy is more often a consequence of other conditions than oxygen shortage during labor, which is what CTG detects. Yet, CTG decreases the risk of neonatal seizures by 50 %, which is indeed associated with complications in the newborn. Another shortcoming of CTG is the high disagreement in interpretation between different experienced labor personnel.

My thesis aims to investigate different aspects of CTG more thoroughly:

In my first study I wanted to elucidate if a more extensive CTG education could lead to better agreement in interpretation. I compared two different labor units in two different parts of Sweden with two different strategies for learning CTG. The conclusion after the study was that extended CTG education might lead to better agreement in CTG interpretation, but more importantly both labor units performed better than expected in terms of CTG interpretation agreement.

In the second study we have collaborated with the Royal Institute of Technology, KTH, and developed a computerized tool for analysis of the CTG trace. We have focused on sudden drops, decelerations, in the fetal pulse rate from its baseline. We wanted to assess the precision of the computer in finding the decelerations and measuring them with regard to duration, depth and area of decelerations. We found evidence that the computer was a valid instrument for this task.

In the third study we continued our focus on the deceleration area, which is a concept not yet used in clinical practice but with possibly great potential to find distressed fetuses. We used 502 CTG tracings for analysis and measured duration, depth and area of decelerations and compared the measurements to fetal lactate concentration. Lactate is produced during anaerobic (oxygen poor) conditions and can be analyzed during labor by taking a small

capillary blood sample from the fetal scalp. We found that deceleration area and duration were more strongly correlated to higher lactate concentrations than was deceleration depth.

In my last study we wanted to explore the role of CTG at admission. In Sweden all women with signs of labor undergo a 20–30-minute admission CTG upon arrival to the labor ward. In some countries, for example the UK and Norway, the use of admission CTG is discouraged in low-risk pregnancies. On the other hand, it is recommended to auscultate the fetal pulse for one minute. Our hypothesis was that among the presumed low-risk women some high-risk fetuses are hidden. For example, as unknown growth restricted fetuses with a malfunctioning placenta, which is the port of oxygen and nutritions to the fetus. We used a large register-based material of > 100 000 low-risk deliveries. We found that among the fetuses with abnormal admission CTG the proportion of small fetuses was twice as high compared to those with normal admission CTG. Moreover, these small fetuses, with abnormal admission CTG, were at a higher risk of severe complications during delivery.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Kardiotokografi (CTG) utvecklades på 1960-talet av bland annat Edward Hon i hopp om att förebygga allvarlig syrebrist hos barnet under förlossningen. CTG mäter fostrets puls och värkarnas kontraktioner kontinuerligt och registrerar dessa som två intilliggande kurvor. Vid tecken på hotande syrebrist kan man på så sätt ingripa och - när så behövs - påskynda den vaginala förlossningen med sugklocka eller utföra ett kejsarsnitt. Dessa operativa ingrepp är förknippade med ökad risk för komplikationer hos både moder och barn. CTG-metoden har flera tillkortakommanden varav en är en hög andel av falskt positiva fall, det vill säga en betydande andel av CTG är icke-normala under förlossningen, trots att barnet är helt välmående. Man har också visat att användandet av CTG medför en ökad risk för kejsarsnitt.

Död eller cerebral pares (CP) är allvarliga komplikationer som kan uppstå i samband med svår syrebrist hos barnet under förlossningen. Vid tidigare försök att studera effekten av CTG har det varit svårt att bevisa en minskad risk för dessa komplikationer. Endast en liten del (10-20%) av alla fall av CP är relaterade till förlossningen och både CP och död är väldigt ovanligt i samband med förlossning vilket medför att det är svårstuderat.

Vad man däremot har kunnat visa är att CTG halverar risken för nyföddhetskramper jämfört med avlyssning av fostrets hjärtljud med tratt eller en doptone – en mindre bärbar hjärtljudsmonitor. Nyföddhetskramper i sig är relaterade till syrebrist och en ökad risk för sjukdom hos barnet. Man har också sett att bedömningen av en CTG-registrering skiljer sig åt markant mellan olika läkare och barnmorskor.

I denna avhandling studeras olika aspekter av CTG:

I) I det första delarbetet undersöktes effekten av tillägget av en strukturerad intensifierad CTG-utbildning i jämförelse med en nationell web-baserad CTG-utbildning på två olika kliniker i Sverige med dessa två olika utbildningssätt. En jämförelse gjordes avseende om överensstämmelsen i CTG-tolkning mellan olika förlossningsläkare var bättre på kliniken med den mer intensifierade utbildningen. Det visade sig att överensstämmelsen generellt gällande CTG-tolkning var mycket högre än förväntat på båda klinikerna och att tillägget av den strukturerade utbildningen verkade innebära en ytterligare något högre samstämmighet i tolkningen.

II) Den andra delstudien genomfördes som ett samarbete med Kungliga Tekniska Högskolan, KTH, där en gemensamt utarbetad dataalgoritm för analys av CTG utvärderades. Både avseende hur väl datorn kunde hitta decelerationer, som är tillfälliga nedgångar i fostrets puls som ofta uppstår i samband med värk, samt hur bra algoritmen kunde mäta bredden, djupet och arean på dessa decelerationer. Det visade sig att datorn fungerade väl för denna uppgift.

III) Även det tredje delarbetet fokuserade på decelerationsarean. Tidigare har man i klinisk vardag koncentrerat sig på hur ofta decelerationer förekommer, hur djupa och framför allt hur breda de är, där bredden är ett mått på hur länge de pågår. Dock verkar just decelerationsarea vara ett relativt bra mått för att förutsäga risken för lågt pH hos nyfödda barn. Här studerades

hur väl ackumulerad bredd, djup och area av decelerationer var associerade med förhöjt laktatvärde i skalpprov (blodprov från fostrets skalp) under förlossningen. Laktat, det vill säga mjölksyra, bildas under syrefattiga förhållanden i vävnaderna. Under förlossningen kan man, vid avvikande CTG, med hjälp av skalpprov analysera laktatkoncentrationen som ett komplement i bedömningen av fostrets tillstånd. Resultaten visade här att en summerad decelerationsarea och bredd under 30 och 60 minuter var bättre på att förutsäga förhöjt laktatvärde jämfört med summerat decelerationsdjup under samma tidsperiod.

IV) I det sista delarbetet undersöktes intagnings-CTG noggrannare. Under 80-talet utarbetades metoden att registrera fostrets hjärtljud med CTG under 20-30 minuter direkt när den födande kvinnan ankommer till förlossningsavdelningen, ett så kallat intagnings-CTG. Målet var att upptäcka de foster som redan hade en begynnande syrebrist tidigt under förlossningsarbetet och sannolikt inte skulle klara av en vaginal förlossning utan risk för framtida men. Metoden har dock kritiserats och flera länder använder inte intagnings-CTG vid lågriskförlossningar. I stället avlyssnar man fostrets hjärtljud med tratt eller doptone under ca en minut efter en värk. Hypotesen här var att det i lågriskgruppen döljer sig en grupp högriskfoster, till exempel till följd av en icke känd tillväxthämning. Över 100 000 lågriskförlossningar undersöktes med hjälp av ett större lokalt register i region Stockholm-Gotland och resultaten visar att de foster som hade ett icke-normalt intagnings-CTG oftare var små för graviditetsåldern jämfört med fostren med normalt CTG. Dessa små foster med icke-normalt intagnings-CTG hade också en påtagligt högre risk för allvarliga komplikationer relaterade till syrebrist.

ABSTRACT

Background: Cardiotocography (CTG) is a tool to assess fetal well-being during labor and to detect early signs of fetal distress and thereby enable timely interventions to reduce neonatal morbidity and mortality. CTG is associated with shortcomings; poor reliability in interpretation, low specificity with a high proportion of false positive tracings indicating fetal distress when not accurate, no proven effect on rare severe outcomes such as mortality and cerebral palsy, but rather contributing to an increased risk of operative delivery. The aims of this thesis was to determine I) if an extended CTG education could lead to better reliability in interpretation compared to a national standard education, II) if a computerized algorithm could be developed with precision in detecting and quantitating decelerations on CTG, III) if deceleration area was a better predictor of fetal acidemia during labor than deceleration depth and duration, IV) the proportion of fetuses with undetected small for gestational age (SGA) in a low-risk population, comparing women that present with normal CTG at admission to labor (admCTG) to those with abnormal admCTG and to compare neonatal outcomes in the two groups stratified on SGA or non-SGA.

Material and methods: The CTG tracings used in paper I-III were extracted from a previous cohort of women in labor, from Karolinska University Hospital, Sweden. All women had undergone fetal blood sampling (FBS) during labor due to suspicious CTG patterns. Six obstetricians from two different hospitals were used as observers in paper I. Inter- and intra-observer reliability using Cohen's and Fleiss kappa was determined for different parameters assessed on CTG. In paper II two obstetricians visually analyzed CTG tracings with variable decelerations and specified duration, depth and area for each deceleration. The computerized algorithm analyzed and quantified the same CTG traces and was compared to the observers using intra-class correlation and Bland-Altman analysis. In paper III the predictive value of deceleration area, duration, and depth for fetal acidemia, measured as lactate concentration at FBS, was explored using receiver operating characteristics, area under curve (ROC AUC). In paper IV, a register-based study, the risk of SGA in relation to the result of admCTG, normal vs abnormal was assessed in low-risk pregnancies. Neonatal outcomes were also determined by multiple logistic regression analysis.

Results: I) The inter- and intra-observer reliability was moderate to excellent at both departments, kappa 0.41-0.93. The department with extended education reached significantly higher interobserver agreement for two of six CTG parameters assessed. II) Computerized assessment of decelerations on CTG compared to visual observers reached excellent intraclass correlation (0.89-0.95) and low bias in Bland-Altman analysis, comparable to that between the two observers. III) The deceleration measures with the best prediction of fetal acidemia was cumulative deceleration area and duration, ROC AUC 0.682 and 0.683 respectively compared to deceleration depth 0.631. IV) The proportion of SGA was two-fold higher among neonates presenting with abnormal admCTG (18.6%) compared to normal admCTG (9.7%). The risk of composite severe adverse neonatal complications was substantially higher in the group with abnormal admCTG/SGA compared to normal admCTG/non-SGA, adjusted odds ratio 23.7 (95% confidence interval 9.8-57.3) **Conclusion:** Inter- and intra-observer agreement was better than expected at both departments studied and extended education might have an impact on interpretation reliability. A novel computerized algorithm for CTG assessment has high precision in detecting and quantifying decelerations. Cumulative deceleration area and duration are better predictors of fetal acidemia than deceleration depth. In presumed low-risk pregnancies there is a group of undetected SGA fetuses that more often present with abnormal admCTG and are at higher risks of neonatal complications.

LIST OF SCIENTIFIC PAPERS

I. 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 obstetricia et gynecologica Scandinavica. 2017;96(4):496-502.

II. Gyllencreutz E, Lu K, Lindecrantz K, Lindqvist PG, Nordstrom L, Holzmann M, et al. Validation of a computerized algorithm to quantify fetal heart rate deceleration area. Acta obstetricia et gynecologica Scandinavica. 2018;97(9):1137-47.

III. Gyllencreutz E, Varli IH, Lindqvist PG, Holzmann M. Variable deceleration features and intrapartum fetal acidemia - The role of deceleration area. Eur J Obstet Gynecol Reprod Biol. 2021 Dec; 267:192-197. doi: 10.1016/j.ejogrb.2021.11.009. Epub 2021 Nov 14. PMID: 34826666.

IV. Gyllencreutz E, Hulthén Varli I, Johansson K, Lindqvist P G, Holzmann M. Admission cardiotocography in low-risk pregnancies and neonatal outcomes, the significance of undetected small-for-gestational age fetuses: a register-based study. Submitted.

CONTENTS

LIST OF ABBREVIATIONS

AGA	appropriate for gestational age
aOR	adjusted odds ratio
ATP	adenosine triphosphate
AUC	area under curve
BF	baseline frequency
bpm	beats per minute
CI	confidence interval
CTG	cardiotocography
FBS	fetal blood sampling
FGR	fetal growth restriction
FIGO	The International Federation of Gynecology and Obstetrics
HIE	hypoxic ischemic encephalopathy
IA	intermittent auscultation
LGA	large for gestational age
min	minutes
NICE	the National Institute for Clinical Excellence
NICHD	the National Institute of Child Health and Human Development
Obs	Observer
OR	odds ratio
RCOG	Royal College of Obstetricians and Gynaecologists
RCT	randomized controlled trial
ROC	receiver operating characteristics
SD	standard deviation
sec	seconds
SFOG	Swedish Society of Obstetrics and Gynecology
SGA	small for gestational age
vs	versus
WHO	World Health Organization

1 INTRODUCTION

In the eternal circle of life, babies have been born and will be born for several centuries and millennia passed and coming. The process of bearing and thereafter delivering the child into the world is challenging and, often forgotten in our part of the world, dangerous both for the mother and the newborn, which women in the developing countries are well aware of and women in Sweden only a century or two ago. Much has been developed to improve the safety of this strenuous process; maternity care, giving birth in health care centers and medical improvements. Nowadays, we expect the woman and baby to come through the childbearing process healthy. Nevertheless, the risks are still there. For example, the force of the uterine contractions can be detrimental to a small group of fetuses who do not have the reserves to cope with the hypoxic nature of labor, risking the life or the future health of that individual.

Cardiotocography (CTG) was developed by Hon in the 50's as a means to assess the well-being of the fetus during labor, with the possibility to intervene in time if signs of distress, i.e., hypoxia emerged. The technique is simple; it includes assessment of the fetal pulse rate continuously as well as the uterine contractions, displayed on screen or paper.

As a matter of fact, there has not been great development of the technique since the 60's, even though attempts have been made to enhance and develop the method. For example, neither supplementing assessment of ST-wave analysis of the fetal echocardiogram or computerized analysis of the CTG have yet been proven to add any benefit to the technique.

Through the years the method as such has been persistently criticized, due to low positive predictive value, low reproducibility in interpretation and difficulties to prove improved neonatal survival or reduced cerebral palsy by its use. However, the latter is more often attributable to brain damage caused earlier in the fetal life in utero than to hypoxic consequences of labor. Also, we are moving through eras of trends where women and labor personnel ask for more or less medical interventions and surveillance. Still, after 60 years of practice, no other technique has yet been proven to excel CTG in detecting ongoing fetal hypoxia.

2 LITERATURE REVIEW

2.1 HYPOXIC PROCESS OF BIRTH

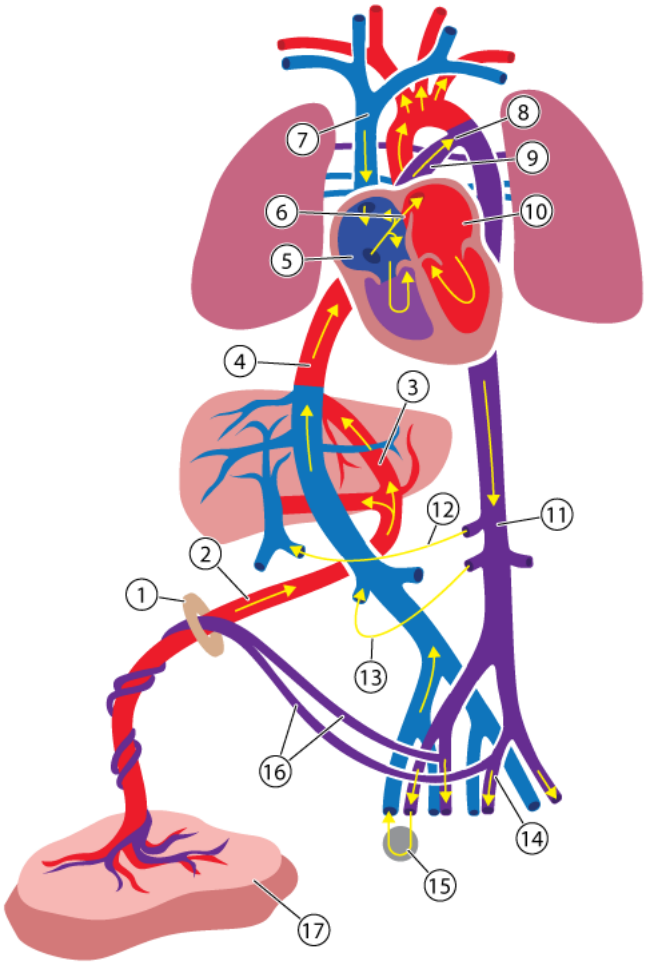
Labor is a hypoxic process for all fetuses. The fetus is designed for impressive resilience to an oxygen-lacking environment in several ways. It has abilities to redirect the blood flow from certain organs (muscles, skin) to other more prioritized ones (brain, heart, adrenal glands) through shunts and vasoconstriction. It has a specially designed hemoglobin with effective binding and release of oxygen.[1] However, the hypoxic process can indeed overcome a threshold where a normal expected hypoxia turns into something more devastating; metabolic acidemia. In this situation the fetus can no longer compensate for the hypoxia and in the end a destroying process of cell death including brain damage occurs.

Placental insufficiency is when the function of the placenta as a nutritional and oxygen port to the fetus is reduced, often due to vascular disease such as preeclampsia, hypertension, and diabetes. In these situations, the fetus can suffer from hypoxia even before the onset of labor, i.e., chronic hypoxia.[2]

During the labor process the uterus contracts regularly with increasing intensity as the labor proceeds. Moreover, the contractions usually become more frequent with advancing stages of labor. During contractions the blood flow to the uterus is decreased.[3] Also, compression of the umbilical cord can occur. As a result, the oxygen supply is intermittently impaired, and hypoxic blood reaches the fetus. However, most fetuses can handle these intermittent hypoxic events without jeopardizing their health.

2.2 FETAL CIRCULATION WITH SHUNTS

Oxygenated blood from the placenta reaches the fetus through the vein in the umbilical cord, which empties into the fetal vena cava inferior through the open ductus venosus (first shunt). The oxygen-rich blood is directed from the right side of the heart to the left one via an open foramen ovale (second shunt). Ductus arteriosus (third shunt) connects the truncus pulmonalis to the descending aorta since there is restricted need of blood supply to the not yet functioning lungs.[4] After oxygen delivery in the different tissues, deoxygenated blood leaves the fetus via the two umbilical arteries and reach the placenta where the gas and nutritional exchange between mother and fetus occurs and the fetal blood is reoxygenated. (Figure 1)



1. Umbilicus
2. Umbilical vein
3. Ductus venosus
4. Inferior vena cava
5. Right atrium
6. Foramen ovale
7. Superior vena cava
8. Ductus arteriosus
9. Truncus pulmonalis
10. Left atrium
11. Aorta
12. Intestines
13. Kidneys
14. Internal iliac artery
15. Lower extremities
16. Umbilical arteries
17. Placenta

Figure 1. The fetal circulation with shunts. From ctgutbildning.se/Löf

2.2.1 Fetal asphyxia, the concepts of respiratory versus (vs) metabolic acidosis

The term asphyxia comes from the Greek language meaning pulseless. The term has thereafter been defined as "a condition of impaired blood gas exchange leading to progressive hypoxemia and hypercapnia with a significant metabolic acidosis".[5]

As a consequence of the physiologic nature of labor with intermittently reduced gas-exchange over the placenta during contractions, oxygen in the fetal blood will decrease, hypoxemia and carbon dioxide may accumulate in the fetal blood, hypercapnia. This leads to respiratory acidosis; the fetus cannot get rid of extra carbon dioxide and carbonic acid is produced. This is indeed common during labor and many neonates are born with a decrease in pH and hypercapnia with no evidence of compromise and an unaffected good prognosis regarding future health.[6] However, severe hypoxia causes anaerobic metabolism in the fetal tissues, leading to metabolic acidosis. This is potentially dangerous in contrast to respiratory acidosis. During anaerobic metabolism the fetus will use glycogen/glucose for energy production which will generate 2 ATP, instead of 36 ATP (aerobic metabolism through citric acid cycle), and lactate acid as a by-product. pH will decrease if the placental gas exchange is not restored, and meanwhile base deficit (BD) will increase since bases are used for buffering the hydrogen ions. Umbilical artery pH < 7.00 at birth is associated to neonatal morbidity and mortality,[7] but levels of pH of < 7.05 and < 7.10 are also correlated to short-term adverse neonatal outcomes and therefore suggested as other reasonable cut-offs for acidemia.[8,9] BD levels of ≥ 16 are associated with adverse neonatal outcomes such as death and severe morbidity.[10] However, levels of BD above 10-12 are correlated with complications in the neonatal nervous system and respiratory system.[11,12] The international cerebral palsy task force uses a cut-off of pH < 7.00 and BD ≥ 12 in umbilical artery cord or early neonatal blood sample as criteria to define an acute intrapartum hypoxic event. Many studies use pH < 7.05 and BD ≥ 12 as definition for metabolic acidemia.[13-15] The worsening acidemia leads to oxidative stress, and inflammation and in the end cell death occur. The fetus will prioritize blood flow to crucial organs such as brain, heart, and adrenal glands by increasing blood flow through the shunts (ductus venosus and ductus arteriosus) and by vasoconstriction to other organs. Nonetheless, in case of severe acidemia, even the central organs will be affected in the end – the heart will fail to keep up the blood pressure and eventually even the pulse rate, and cell death and damage develop in the brain.[16] Hypoxic ischemic encephalopathy (HIE), is a feared consequence of fetal asphyxia with short- and long-term complications such as cerebral palsy, epilepsy, and cognitive impairment.[17] The incidence of HIE grade 2–3 in term pregnancies in Sweden between 2016-2020 was 7 per 10 000 term deliveries. With an Apgar score < 4 at 5 minutes, the risk of HIE was approximately 19%. Corresponding figure for Apgar scores < 7 at 5 minutes was 5%.[18]

2.3 FETAL GROWTH RESTRICTION (FGR)

Some fetuses will be subjected to an unfavorable intrauterine environment and their full growth potential will not be reached. The most common cause of FGR is placental insufficiency, depriving the fetus from the nutrients and oxygen required for uncompromised development and growth. These fetuses will spare their energy by a reduced growth capacity, reduced fetal reserves such as glycogen depots, and reduced fetal movements.

Placental insufficiency is a predominantly vascular disease, either associated with gestational hypertensive disease or pre-existing maternal morbidity, such as hypertension, diabetes, autoimmune diseases and heart- and lung diseases. Other conditions associated with FGR are intrauterine infections and chromosomal abnormalities. Tobacco use and the use of other substances such as alcohol and cocaine also contribute to increased risks of FGR.[19] Fetuses affected by FGR have increased risks of morbidity and mortality.[20]

Fetal growth during pregnancy is estimated by ultrasonographic measurement of fetal biparietal diameter, abdominal circumference or mean diameter and femur length. When the fetus is eventually born the terms small for gestational age (SGA), appropriate for gestational age (AGA) and large for gestational age (LGA) are introduced, with original cutoffs at < 10^{th} percentile, 10^{th}-90^{th} percentile and > 90^{th} percentile respectively.[21] However, SGA newborns are not by definition FGR, since some newborns will be genetically small and not exposed to growth restriction. Moreover, some AGA newborns have suffered from growth restriction, not reaching their full growth potential. Hence, repeated ultrasound measurements, assessing growth velocity, are needed to identify true FGR. Another indication of FGR detected by ultrasound are the presence of abnormal Doppler velocimetry of fetal blood vessels.[22]

The threshold for SGA differs in different countries and literature, but the 10^{th} percentile cutoff has been the most used.[23,24] Studies report that the neonates with birth weight < 3^{rd} percentile are at the highest risk of mortality and morbidity even though risks start to increase from higher percentiles.[20,25] A consensus definition of FGR was elaborated 2016, defining FGR as either estimated fetal weight of 3^{rd} percentile for gestational age or 10^{th} percentile in combination with abnormal Doppler velocimetry of fetal blood vessels.[26]

To detect this group of fetuses, several different strategies have been developed. In Sweden, the woman undergoes regular measurements of fundal height at midwifery led antenatal care. When risk factors for FGR are present or arise during pregnancy, individualized ultrasound weight estimation with or without fetal Doppler velocimetry is performed. However, the less resource demanding measurement of fundal height has proven to be inadequate in detecting FGR fetuses.[27] Moreover, some fetuses will suffer from FGR without known risk factors such as maternal morbidity.

2.4 CTG

2.4.1 Development and technology

Cardiotocography (CTG) was developed by Hon and others in the 1950's and 60's as a means to assess the well-being of the unborn fetus before and during labor.[28] It is a technique that measures the fetal pulse rate either by Doppler ultrasound detection by an external monitor on the maternal abdominal wall or by gauging the R-R interval on the fetal echocardiogram by a needle electrode attached to the presenting part of the fetus, which means that the fetal membranes need to be ruptured. The fetal pulse rate is then continuously presented on paper or screen with different speeds depending on geographic location; for example, 1 cm/min in Europe and 3 cm/min in the United States.

The CTG also measures the contractions of the uterus by a transducer attached to the maternal abdominal wall, at the fundus uteri, and is likewise displayed continuously on the cardiotocogram below the fetal pulse rate. The onset, the peak, the duration and the frequency of the uterine contractions are assessed (Figure 2).

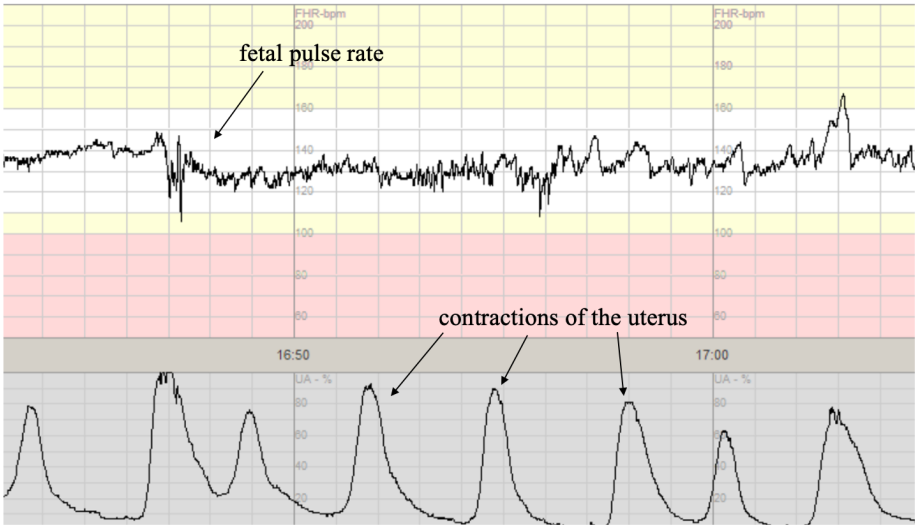


Figure 2. CTG assessing fetal pulse rate and uterine contractions, paper speed 10 mm/min, from ctgutbildning.se/Löf

2.4.2 Variables of interest

2.4.2.1 Baseline frequency

Baseline frequency (BF) is defined as the mean pulse rate during a 10-minute period without including accelerations or decelerations. A normal BF lies between 110-160 beats per minute (bpm) during labor. In Sweden, a new CTG guideline was launched in 2017 by the Swedish Society of Obstetrics and Gynecology (SFOG).[29] The previous national guideline had a limit of 150 bpm for tachycardia, but to harmonize with the revised international classification systems by The International Federation of Gynecology and Obstetrics (FIGO)[30] the limit

was changed to 160 bpm. [29] A BF of less than 110 bpm is considered bradycardia. Nonetheless, there is recent evidence that a BF between 150 and 160 bpm at or after 40 gestational weeks is associated with neonatal acidemia compared to 110 to 149 bpm.[31] It is known that BF decreases with progressive gestational age.[32]

2.4.2.2 Variability

Heart rate variability is the occurrence of small changes/oscillations in BF over short periods of time; this is visualized as an irregularity in the BF line. Variability is the bandwidth of the BF assessed during one minute. A normal variability ranges from 5-25 bpm, a reduced variability is between 2-5 bpm, and an increased variability >25 bpm. Absent variability is defined as oscillations below 2 bpm, a pattern requiring prompt action. However, there are questions whether an isolated reduced variability without other CTG abnormalities is correlated to fetal hypoxia. In a study by Holzmann et al, when assessing isolated reduced variability, the proportion of fetuses with an increased lactate concentration (> 4.8 mmol/L) in a fetal blood sample (FBS) during labor was comparable to a group of fetuses with normal baseline frequency and normal variability (2.6 vs 2.5%).[33] Cahill et al showed that minimal variability was not associated with neonatal acidemia (pH < 7.10) compared to normal variability, OR 0.68 (95% CI 0.27-1.71).[34] Triebwasser used the same cohort as Cahill, but analyzed the fetuses undergoing operative delivery during the second stage and found that minimal variability was as common in the group with pH < 7.2 (7.1%) compared to the group with pH >7.2 (7.3%).[35]

2.4.2.3 Accelerations

Accelerations are occurrences of increasing periodic pulse rates, lasting more than 15 seconds and reaching a pulse rate more than 15 bpm higher than the BF. Accelerations are an assuring sign of a well-oxygenated fetus. When provoking accelerations by different fetal stimulation tests during labor a negative test result (occurrence of accelerations) was correlated to a very low likelihood ratio of fetal acidemia, as low as 0.06 for digital stimulation.[36] It is not uncommon that accelerations disappear during labor. However, this is most likely not associated with development of fetal acidemia. When studying the time that had passed since the last acceleration and correlation to acidemia at FBS, there was no significant difference in proportion of acidemia between fetuses with accelerations occurring during the last 20 minutes compared to those with no acceleration in the last 60 minutes.[37]

2.4.2.4 Decelerations

Sudden drops in the fetal pulse rate, decelerations, are very common during labor. Analogous to accelerations, decelerations are negative deviations from baseline that last for more than 15 seconds and reach at least 15 bpm lower than BF. Decelerations can have varying shape and temporal relationship to the uterine contractions. During the years, there have been numerous theories behind deceleration origin and pathology. Some decelerations are viewed as harmless and others as potentially threatening to the fetus.

Uniform decelerations are characterized by a similar feature from one another and are subdivided into uniform early and uniform late decelerations.[38] They are internationally titled early and late decelerations respectively, and these terms will be used in this thesis. Early decelerations are mirroring the contraction, with nadir at the same time point as the peak of the contraction. These are considered as harmless and subsequently therefore classified as normal in the Swedish guidelines for intrapartum CTG. The theory behind early decelerations, is that a pressure on the fetal head leads to a vagal response and a corresponding decrease in the fetal pulse rate.

On the other hand, late decelerations are skewed in their onset and nadir in relation to the uterine contraction; the deceleration starts after the peak of the contraction and returns to baseline after the contraction has ceased (Figure 3). The theory behind late decelerations, which are skewed in their onset in relation to the contraction, have been that they are chemoreceptor mediated as a response to fetal hypoxemia.[30] Due to placental insufficiency, there is a reduced blood flow to the placenta; during contractions hypoxic blood is pooled and re-perfused to the fetus after the contraction has peaked, leading to chemoreceptor activation and hence decelerations. Shallow late decelerations with an amplitude of 10-15 bpm can occur and are often accompanied with reduced variability (see Figure 6 below).

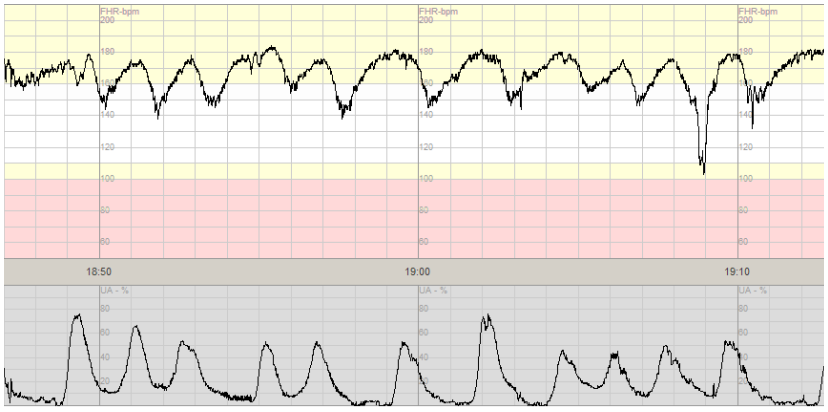


Figure 3. Late decelerations, from ctgutbildning.se/Löf

Variable decelerations are characterized by a sudden drop in the fetal pulse rate, reaching nadir in < 30 seconds and with shifting appearance from deceleration to deceleration (Figure 4). In Sweden, the variable decelerations are further subdivided into complicated and un-complicated; variable complicated decelerations last more than 60 seconds in duration in contrast to the un-complicated variable deceleration. Variable decelerations are presumed to be less harmful than the late decelerations unless the baseline frequency or variability is abnormal.

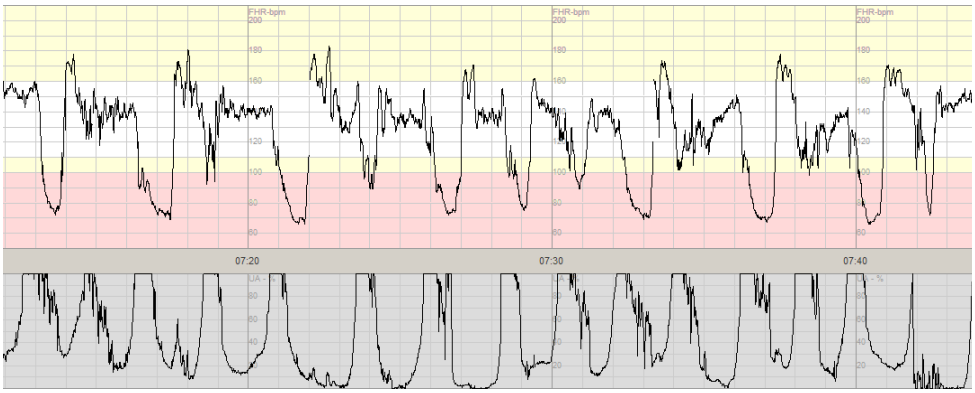


Figure 4. Variable decelerations. From ctgutbildning.se/Löf

Historically, the prevailing hypothesis of variable decelerations has been that they are baroreceptor mediated. Compression of the umbilical cord raises the arterial pressure due to increased peripheral resistance. This leads to activation of baroreceptors and a subsequent rapid drop in the fetal pulse rate.[39,40]

Increasing evidence, from animal studies, during recent years have questioned the concept of baroreceptor-mediated variable decelerations.[41] Indeed, it seems that even the variable decelerations could be caused by chemoreceptor activation due to a fall in oxygen tension due to uterine contractions, irrespective of whether cord occlusion is present or not. The uterine contraction leads to decreased blood flow over the placenta to the fetus, causing repeated brief periods of hypoxemia. The chemoreceptor activation leads to a rapid, parasympathetically mediated, decrease in fetal heart rate, as well as a sympathetically mediated peripheral vasoconstriction.[42] The fall in fetal heart rate is a protective response, leading to reduced myocardial work. The vasoconstriction on the other hand sustains or increases mean arterial blood pressure and centralizes blood flow to central crucial organs such as brain, heart, and adrenal glands. Animal studies, performed on sheep fetuses, have shown that occlusions of the umbilical cord generate repetitive variable decelerations but are tolerated among vigorous fetuses as long as the occlusion is brief and in low frequency; 1:5 (1 minute occlusion every 5 minutes).[43] During these attempts the sheep fetus copes with the recurrent decrease in blood flow for hours. However, when the occlusion is prolonged in duration to 2:5 minutes or the interval between occlusions are shortened 1:2.5 minutes, even the vigorous fetuses will develop severe acidemia as time proceeds. The decelerations occurring will be deeper and longer in duration as acidemia develops.[44,45] When sheep fetuses had been subjected to chronic hypoxia mimicking uteroplacental deficiency their ability to endure 1:5 minutes occlusion was abolished, leading to EEG suppression and seizure activity.[46]

Decelerations in international context

In the 1970's Krebs coined the expression atypical decelerations.[47] Atypical decelerations have one of the following features: a) loss of initial acceleration, b) slow return to baseline frequency, c) loss of secondary acceleration, d) prolonged secondary acceleration, e) biphasic

decelerations, f) loss of variability during deceleration, g) continuation of baseline at a lower level after the deceleration. The feature associated with the highest percentage of low Apgar scores at 1 and 5 minutes according to the study by Krebs were loss of variability during deceleration.[47]

The FIGO classification also subgroups decelerations into early, variable, late and prolonged.[30] The variable decelerations are V-shaped and reach nadir in 30 seconds' time and have a good variability in the deceleration. Late decelerations are U-shaped and/or with reduced variability. According to the guidelines, if there is a good registration of contractions, the deceleration starts at least 20 seconds after the onset of the contraction and reaches nadir after the peak of the contraction. Hence, this classification of decelerations differs from the Swedish definition, where several variable complicated decelerations will be classified as late, due to a U-shape.

The guidelines in the UK, the National Institute for Clinical Excellence (NICE) and the Royal College of Obstetricians and Gynaecologists (RCOG) also subdivides decelerations as early, variable, late and prolonged. The variable decelerations are regarded as concerning when one or more of following occur: lasting > 60 seconds, reduced variability within the deceleration, failure to return to baseline, biphasic shape or no shouldering.[48]

The National Institute of Child Health and Human Development (NICHD) classification, used in the United States, subdivides decelerations as early, late, variable, or prolonged.[49] Variable decelerations are decelerations with a rapid decline, reaching nadir in 30 seconds. Atypical features of decelerations are considered. When compared to the NICE guidelines and the original Krebs criteria, there is a contradiction in the view of atypical decelerations. Krebs thought that the loss of initial or secondary accelerations, before or after the deceleration was abnormal and hence classified as an atypical deceleration. However, in the NICHD classification, the appearance of shoulders is thought to be an atypical deceleration.

Hence, the nomenclature and definitions of decelerations differs internationally (Table 1). Some authors suggest that to distinguish between late and variable decelerations during labor is of less importance.[42] The true late decelerations will be present at the earliest phase of labor and are indeed signs of a compromised fetus. Late decelerations occurring later in labor are comparable to the variable ones. However, this standpoint has been criticized.[50]

Cahill and co-workers studied atypical decelerations more recently with a publication from 2012 and found no association between atypical features (according to the NICHD classification) of decelerations and fetal acidemia (pH < 7.10).[34] All patients reached full cervical dilation and the CTG trace during the last 30 minutes before delivery were studied. Hence, all women with atypical decelerations earlier in labor and/or delivered by cesarean before complete cervix dilatation were not included in the study. Also, the number of acidemic fetuses was low, and there is a possibility of insufficient power to detect differences in outcome. Hamilton et al studied atypical decelerations during the last 4 hours of labor in three groups of neonates, normal (BD < 8), metabolic acidosis (BD > 12), and abnormal (BD

> 12 and encephalopathy). They observed that prolonged decelerations (> 120 seconds), decelerations with 60's criteria (two of following: > 60 bpm magnitude in beat loss, nadir reaching below 60 bpm, duration of > 60 seconds), and loss of variability in the deceleration were features that differed significantly between the three groups. However, the areas under the curve (AUC) in the receiver operating characteristics (ROC) analysis were quite low.[51] Furthermore, Gamboa et al studied atypical decelerations during the last 30 minutes of birth in a case-control study and found that slow return to baseline and loss of moderate variability within the variable deceleration are patterns associated with acidemia at birth (pH < 7.10 in arterial umbilical cord gas).[52]

Prolonged decelerations

The definition of a prolonged deceleration differs among the different classification systems, with a duration limit of > 60 seconds or up to > 3 minutes.

	Variable decelerations	Late decelerations	Prolonged decelerations	Pathology
NICE/RCOG -17[48]	Variable, intermittent periodic slowing of FHR with rapid onset and recovery.	Uniform, repetitive periodic slowing of fetal heart rate. Onset at mid to end of contraction, nadir >20 sec after peak of contraction. When reduced variability an amplitude <15 bpm is included.	> 60–90 seconds	**Non-reassuring:** Variable decelerations with no concerning characteristic* > 90 minutes. Variable decelerations < 50% of contractions with any concerning characteristic* > 30 min. Variable decelerations > 50% of contractions with any concerning characteristic* < 30 min. Late repetitive decelerations < 30 min without bleeding or meconium. Single prolonged < 3 minutes. **Abnormal:** Repetitive (> 50% of contractions) variable decelerations with any concerning characteristic* >30 min. Late decelerations > 30 min. Single prolonged deceleration > 3 minutes. * concerning characteristics: lasting > 60 seconds, reduced variability within deceleration, biphasic shape, failure to return to baseline, no shouldering
NICHD -08[49]	Abrupt decrease in heart rate. Reaching beginning of nadir < 30 seconds	Gradual decrease > 30 sec to nadir and return to baseline. Nadir after the peak of contraction.	> 2 min	**Category II:** recurrent (> 50% of contractions) variable decelerations, prolonged decelerations, recurrent late decelerations with moderate variability, variable decelerations with slow return to baseline, biphasic shape, tachycardia after deceleration, shoulders, fluctuation in fetal heart rate in trough of the deceleration **Category III:** absent baseline with recurrent late or variable decelerations or bradycardia
FIGO -15[30]	V-shaped, reaching nadir < 30 seconds. Good variability within deceleration. Rapid recovery to baseline.	U-shaped and/or with reduced variability. Gradual onset > 30 seconds to nadir and return to baseline. Start > 20 seconds after onset of contraction. Amplitude 10-15 bpm when no accelerations and reduced variability.	> 3 min	**Pathological**: Repetitive late or prolonged decelerations > 30 min (> 20 min when reduced variability). One prolonged > 5 min.
SFOG – 17[29]	Varying shape, depth and duration between different decelerations. Quick and prominent amplitude. Uncomplicated: <60 seconds. Complicated > 60 seconds.	Start after onset of contraction, reach nadir 30-60 seconds after top of contraction. Return to baseline after end of contraction.	> 2-3 min	**Aberrant:** repetitive (> 50% of contractions) variable complicated decelerations with normal BF and variability. **Pathological:** repetitive late > 30 min (> 20 min when tachycardia or reduced variability), repetitive variable complicated with tachycardia or reduced variability > 20 min. Repetitive prolonged > 3min. Single prolonged > 5 min.

Table 1. Decelerations comparison, definition and pathology, international and Swedish classification

Deceleration area

Deceleration area is a topic that has been of researchers' interest over the years, as far back as during the 1970's. Shelley and Tipton published data in 1971 and reported that with an increasing "dip area" the mean Apgar score and pH was decreasing.[53] The calculation of the dip area was manual and complex.

Today, there is an increasing focus on the area of decelerations and recent studies indicate that deceleration area is among the best predictors of fetal acidemia. This confirms the findings from the animal studies mentioned above where decelerations were longer and deeper with increasing fetal acidemia. Some studies have used the approximation of deceleration area as deceleration width (duration) × depth/2. Cahill and co-workers evaluated deceleration area by visual assessment of 5388 CTG tracings and used this approximation for deceleration area.[34] They found that deceleration area was the best predictor of neonatal acidemia at birth (pH < 7.10), compared to other known patterns associated with acidemia such as isolated tachycardia and repetitive late decelerations. However, few decelerations are indeed V-shaped, but rather U-shaped, which leads to an underestimation of deceleration area. Gamboa et al, studied different NICHD and non-NICHD features of CTG tracings and their relation to fetal acidemia at birth.[54] Of the different patterns, deceleration area (calculated with the formula above) was the best predictor of acidemia. Also, deceleration duration > 60 seconds and decelerations reaching below 60 bpm, called severe decelerations in this study, were significantly more common in the acidemic group compared to the non-acidemic group.

Nevertheless, if deceleration area is to be assessed continuously during labor, a computerized calculation of deceleration area is necessary since the labor personnel cannot accurately assess this by visual examination of CTG tracings in real-time accurately. Two previous studies, using a computerized algorithm for deceleration area calculation, investigated the relation between deceleration area and neonatal acidemia.[55,56] Furukawa et al also used a computerized algorithm to measure deceleration area, which was the best predictor of neonatal metabolic acidemia defined as BD > 12 mmol/L.[57]

Deceleration capacity

Deceleration capacity is another technique of computer-driven analysis of decelerations on CTG, analyzing all falls in fetal pulse rate between two consecutive heartbeats. This is performed via phase-rectified signal averaging (PRSA). The advantage with assessment of deceleration capacity is that determining BF is less important compared to when measuring deceleration area. Deceleration capacity has been associated with fetal acidemia during labor [58] and in fetal sheep models.[59,60] Furthermore, during experimental cord occlusion series, deceleration area and deceleration capacity can predict fetal hypotension in sheep.[59]

Physiology of fetal hypoxia – CTG interpretation

Some experts argue that a more physiologic approach should be used when interpreting CTG to detect fetal hypoxia, rather than strictly using the classification systems. They recommend that the progression of fetal heart rate patterns should be evaluated.[61,62] Signs on CTG of progressive fetal intrapartum compromise or subacute hypoxia, called Hon's pattern or "stair-step to death" described in the 1960's, commence with the occurrence of decelerations that become wider and deeper during progression of hypoxia and is followed by an increment in baseline frequency. The fetus compensates for the intermittent hypoxia caused by uterine contractions by increasing its heart rate, with episodes of tachycardia which eventually become persistent. By progression of hypoxia and developing acidemia, the variability will decrease, followed by absent variability and preterminal bradycardia.[61,63]

2.4.3 Shortcomings

CTG as a method has been criticized because of several different shortcomings. First, critics emphasize that CTG has not been proven to reduce perinatal mortality or cerebral palsy but has rather increased the rate of cesarean delivery and operative vaginal delivery compared to intermittent auscultation (IA).[64] Auscultation is performed by Pinard's stethoscope, a hollow wood or metal horn, or by handheld Doppler fetal monitoring, assessing fetal pulse rate during at least one minute after a contraction. This is repeated every 15 minutes during the first stage of labor and after every contraction or every 5 minutes during the active second stage. However, perinatal mortality and cerebral palsy luckily are rare events. In Sweden, the incidence of perinatal mortality is approximately 5/1000 deliveries (whereof around 1/1000 deliveries is due to early neonatal death during the first week of life) and the incidence of cerebral palsy is approximately 2/1000 live births.[65] The studies on CTG vs IA at hand are under-powered to detect a change in these rare outcomes. Nonetheless, in a high-risk setting in Greece, with a high perinatal mortality rate (~20/1000 deliveries), continuous CTG during labor was associated with a lower mortality compared to IA.[66] Moreover, studies have observed a 50% reduction in neonatal seizures when using CTG compared to IA.[64,67,68] Neonatal seizures are known consequences of fetal asphyxia while cerebral palsy is more often explained by antenatal neurological damage than by birth asphyxia.[69]

2.4.3.1 Reliability

Several studies have investigated the reliability in CTG interpretation. These studies report a wide disagreement in CTG assessment, both between different observers (inter-observer reliability) and in the same observer, when one observer assesses the same CTG tracing with some time in between (intra-observer reliability).[70-72] In a study by Palomäki which investigated the interobserver reliability by calculating proportion of agreement, the agreement was poor especially among the abnormal tracings, and acceptable or good among normal tracings when evaluating different characteristics on CTG. [71] Since proportion of agreement does not take chance agreement into account, several studies have used the kappa statistics to evaluate interobserver agreement, which evaluate agreement beyond chance

agreement. Chauhan studied interobserver reliability among five observers using the kappa statistics and the agreement was generally only poor to fair for different CTG characteristics. Tachycardia was the only variable where the observers performed good agreement.[72] However, Blackwell et al reported a moderate interobserver agreement using the kappa statistics when three observers categorized CTG tracings of different severity according to the NICHD classification.[73] The intraobserver agreement was even better. Epstein et al observed good interobserver agreement in all variables on CTG assessment as well as categorization of the CTG tracing according to the NICHD classification among physicians with different clinical experience.[74]

2.4.3.2 Specificity

Another important shortcoming is the very poor specificity of CTG as a diagnostic tool for hypoxia. The sensitivity is known to be excellent; a normal CTG trace is assuring for a well-oxygenated fetus. The false negative rate is very low. However, the specificity is very low, in other words the false positive rate is high.[64] This entails that the majority of fetuses presenting with abnormal or pathological tracings will be well oxygenated. The false positivity implies an increased risk of unnecessary interventions during labor with operative delivery, e g vacuum extraction or cesarean delivery.[75] In a study by Jackson et al, in 80% of the deliveries a pattern of category II-tracing occured sometime during labor.[76]

2.4.3.3 Wide range of classification systems

Internationally, we do not speak the same language when it comes to CTG nomenclature. After estimating all the different variables on CTG, all classification systems consist of a cross table for classifying the CTG as normal, suspicious, or pathological. Most classification systems consist of a three-tier system (normal, suspicious, pathological, or category I, II, and III), but there are also five-tier systems in use. In Sweden, we had a four-tier system until 2017, thereafter a three-tier system.

Several different classification systems are in use internationally. In 2015, FIGO made an effort to change this, by creating new consensus guidelines involving 46 international CTG experts. [30] In Sweden, the classification system was changed in 2017 to harmonize with the FIGO guideline, but still there are some differences.[29] The American classification system, the NICHD classification, is a three-tier classification which categorizes CTG as category I (normal), category II (indeterminate), and category III (abnormal). The majority of CTG tracings during labor will at some time point be classified as category II, since this category spans over a wide variety of different features of the fetal pulse rate. The guidelines were updated in 2008.[49] In the UK, NICE/RCOG has developed guidelines for CTG interpretation, consisting of a three-tier system that categorize a CTG tracing as normal, suspicious, or pathological.[48]

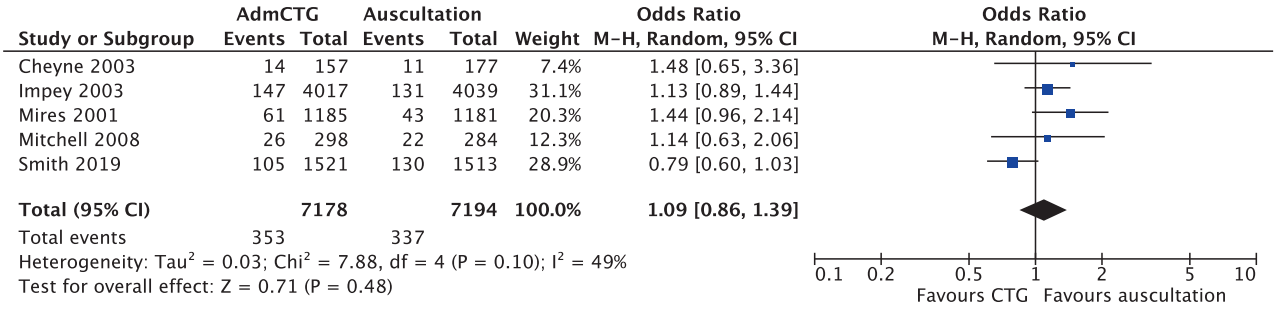
The most obvious change in the new Swedish classification system is that a greater baseline frequency is accepted as normal, up to 160 bpm, compared to 150 bpm earlier. Variable decelerations become pathological first when repetitive and complicated and the baseline

frequency is > 160 bpm or variability is reduced. Hence, more patterns have been normalized with the new classification. The new FIGO 2015 and Swedish 2017 system has been criticized as being too liberal in normalizing patterns, resulting in more neonates subjected to acidemia[77]. A nation-wide study in Sweden comparing neonatal outcomes before and after introduction of the new classification observed a higher proportion of birth acidemia, low Apgar scores, and neonatal seizures after the new classification was launched. However, the risk of operative delivery, both cesarean and vacuum, was lower. Nevertheless, one must be cautious when speculating of causation since other factors could have contributed to the results.[15]

2.4.4 Admission CTG

Admission CTG (admCTG) is performed immediately when arriving at the labor ward with symptoms of labor onset. The procedure was developed in the 1980's after the work of Ingemarsson, aiming to detect fetal distress already present at admission.[78] Some vulnerable fetuses will not cope with the laboring stress, hence giving rise to increasing acidemia and threat of health and life of the fetus. During the years, the necessity of admCTG has been debated and the World Health Organization (WHO), NICE, Norwegian, and Canadian guidelines do not recommend admCTG in low-risk deliveries, but rather recommend auscultation of the fetal pulse rate at admission.[48,79-81] Nonetheless, most cases of fetal asphyxia occur in low-risk pregnancies without antepartum risk factors.[82] Randomized controlled trials (RCTs), studying low risk pregnancies, have not been able to prove any benefit from using admCTG compared to auscultation at admission, rather implying an increased risk of cesarean delivery. In Sweden, the routine of universal admCTG has not changed, but the method has been questioned lately.[83] However, in the largest RCT performed in this field, risk analysis was made after early amniotomy which is not tradition in Sweden.[84] In cases of meconium-stained liquor, women could not be included in the study. The women were also treated with one-to-one midwifery care, unfortunately not yet clinical praxis in the Swedish setting. Furthermore, neonatal outcomes experienced several hours after admCTG have been in focus of interest. Hence, the original idea to detect ongoing fetal hypoxia at admission has fallen into oblivion. However, the risks of admCTG needs to be appointed, when discussing risk-benefit of a procedure. A Cochrane report from 2017 stated that there was a 20% higher risk of cesarean delivery when using admCTG compared to IA, although the results were borderline significant, RR 1.20 (95% CI 1.00-1.44).[85] In 2019, the ADCAR trial, an RCT including 3034 women was published, reporting no higher risk of cesarean delivery, but rather a possible beneficial effect of admCTG: 6.9% cesarean sections vs 8.6% in the IA group.[86] When I add the ADCAR study, Smith 2019, to the Cochrane meta-analysis the association between admCTG and cesarean delivery is abolished (Figure 5). Nevertheless, the use of admCTG increases the risk of continuous CTG and maybe even the use of FBS.[85]

Figure 5. Comparison admCTG vs auscultation (low-risk women), outcome cesarean delivery



The challenge with proving the effect of admCTG is that the proportion of fetuses at risk, i.e., the as of yet undetected vulnerable fetuses, for example due to undetected growth restriction, are indeed rare in the low-risk population. Very large RCTs are thus needed to detect a statistically significant effect of admCTG. Indeed, the Cochrane report on admCTG vs IA states that "no trial or meta-analysis will be adequately powered to detect differences in perinatal mortality".[85] The number needed to treat is probably large and therefore the question must be if we should abandon the routine with admCTG, when it might be of great importance to a small proportion of fetuses who cannot cope with the stress of labor? The growth restricted fetus more often presents with discreet late decelerations at an early stage of labor, a fetal heart rate pattern most likely not detectable with auscultation alone.[87] An example of this is seen in figure 6 below which shows the admCTG of a fetus born SGA 10th percentile.

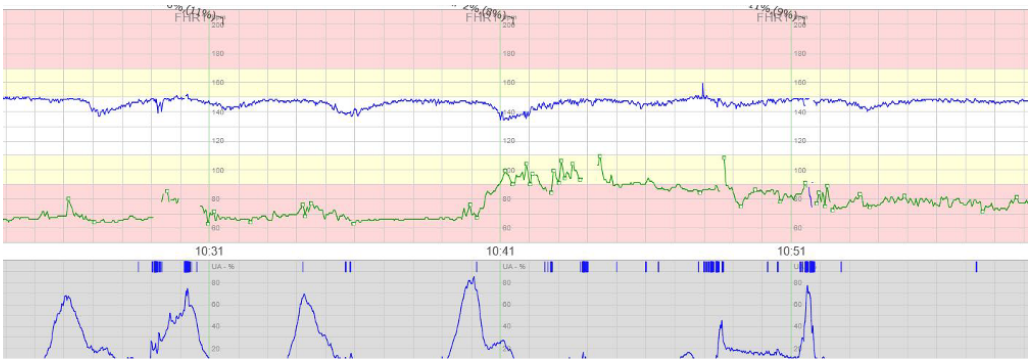


Figure 6. AdmCTG of a fetus born SGA 10th percentile in a presumed low risk pregnancy. Note the normal BF, reduced variability, and discreet late decelerations. The concern is whether this pathological pattern could have been detected by auscultation.

2.4.5 Computerized CTG

Efforts have been made in computerizing the assessment of CTG and different systems have been developed. The antenatally used short-term variation developed by Dawes and Redman [88] has proved of use antenatally[89] especially in growth-restricted fetuses.[90] In terms of

computerized CTG during labor, the hope has been that the computer could overcome the inevitably human factor in CTG interpretation; the challenge with interobserver variability and human errors in interpretation during stress and fatigue. Two of these systems for intrapartum use have been scrutinized by RCT studies.[14,91] One was conducted as a multi-center study in the UK and the computerized system analyzes the CTG and provides real-time visual color and sound alerts. The system detects features that are correlated to hypoxia such as repetitive decelerations. More than 7000 women were recruited in the study, based on a sample size calculation of a hypothesized reduction in metabolic acidosis (pH < 7.05 and BD ≥ 12) from 2.8% to 1.8%. Nevertheless, there were no differences in primary outcome (metabolic acidosis) or secondary outcome (e.g., cesarean section rate, HIE etc.) between the intervention and control arm. However, the incidence of metabolic acidosis was far lower than expected and the power might have been insufficient to detect differences in outcome.[91] The second study, the INFANT trial, enrolled > 46 000 women. The INFANT system is a decision-supported computerized system that analyzes the CTG trace continuously and signals through a color-coded alert system. The main outcome studied was a composite of poor neonatal outcome and developmental progress at two years of age. Despite the large sample size there was no evidence of any benefit of the INFANT system compared to CTG alone, regarding all outcomes studied. [14] These studies as well as another RCT of a smaller size (720 women included), [92] have been further investigated by a systematic review[93] with no differences in primary (metabolic acidosis) or secondary outcomes (mode of delivery, admission to neonatal intensive care unit, HIE and death).

A computerized system prototype that uses deceleration capacity and/or the presence of a non-reactive initial trace, adjusted for preeclampsia and thick meconium was investigated. The comparison was to detect fetal compromise in a retrospective cohort using a computerized alarm system, based on the algorithm above, compared to clinical decision to intervene, i.e., operative delivery due to fetal compromise. Sensitivity and false-positive rate were evaluated with trends toward better sensitivity and reducing the risk of false-positive results.[58]

Machine learning techniques are indeed very interesting in the field of CTG, but not yet in clinical practice.[94,95] Cömert et al described the combination of four different algorithms for machine learning examining 552 CTG recordings to predict neonatal acidemia, defined as pH < 7.20 at birth. They demonstrated a sensitivity of 77.4%, a specificity of 93.9% and ROC-AUC of at most 0.887 with the use of machine learning.[95]

2.5 ADJUNCT TECHNOLOGIES

To overcome the problem with low specificity of CTG, efforts have been made to develop adjunct technologies, to decrease the risk of unnecessary operative interventions during birth.

2.5.1 Fetal blood sampling

FBS during labor was one of the earliest adjunct technologies that was elaborated. This method was introduced as early as the 1960's after research done by Saling on sheep

fetuses.[96] By making a small incision on the fetal scalp a small blood sample can be collected. At first, pH was analyzed, but during the 1980's a new technique with analysis of blood lactate concentration was introduced.[97] Lactate is produced in the cells during anaerobic conditions, i.e., hypoxia. The lactate analysis requires a lesser amount of blood compared to pH analysis and therefore the success rate to perform a sufficient sample is thus higher.[98] Studies have observed that lactate analysis is as safe as pH analysis.[99,100] Also, theoretically, lactate might be more favorable than pH as a method as it specifically measures the product of anaerobic metabolism, i.e. a sign of metabolic acidemia, while pH also decreases in fetuses with respiratory acidosis.[101] However, when using lactate concentration in FBS, different reference values have to be set for different meters, since the lactate concentration differs between different lactate measurement systems. For example, the previously used Lactate Pro™ (Arkray, Kyoto, Japan) had a widely used cut-off value of 4.8 mmol/L for fetal acidemia[102] whereas for the new device Lactate Pro 2™ (Arkray, Kyoto, Japan) a cut-off of 7.3 is suggested.[103] The effectiveness of FBS has been reviewed with indices of reduced cesarean delivery by use of FBS combined with CTG.[104,105] The Flamingo RCT was performed to elucidate this relationship further, comparing FBS in addition to CTG versus CTG alone. However, the number of recruited participants (n= 123) was far lower than planned after power analysis (n= 600). No differences were observed in proportions of cesarean delivery or composite neonatal morbidity.[106]

2.5.2 Pulse oximetry

Pulse oximetry is measured by attaching a probe to the fetal head, cheek, or back during labor, thereby assessing the oxygen saturation in fetal blood. Yet, the method has not been proven to improve fetal outcomes or reduce cesarean sections compared to CTG monitoring alone.[107]

2.5.3 ST wave analysis (STAN)

ST analysis of fetal echocardiogram during labor was developed in Sweden. At first, the method showed promising results,[108] but in a large RCT published 2015, no benefit of the method could be proven compared to CTG alone.[13] Thereafter, a meta-analysis including six randomized trials (n= 26 529) was carried out reporting no additional benefit of STAN in terms of perinatal outcome or cesarean delivery compared to CTG alone.[109]

3 RESEARCH AIMS

The research in this thesis aim is to explore different aspects of CTG and investigate new strategies more thoroughly.

- To determine if addition of a regular extended CTG training could enhance CTG inter- and intra-observer reliability compared to current web-based CTG training alone
- To validate a new computerized CTG algorithm and its performance in accurately identifying decelerations and quantitating their width, depth and area
- To determine if cumulative deceleration area is a better predictor of fetal acidemia during labor compared to current measurements, such as deceleration depth and duration
- To determine if the proportion of undetected SGA fetuses is higher when the fetus presents with abnormal compared to normal admission CTG in a low-risk population. A secondary objective was to explore if the SGA fetuses with abnormal admission CTG are more susceptible to neonatal complications compared to non-SGA fetuses with normal admission CTG

4 MATERIALS AND METHODS

The four studies included in this thesis differ substantially in methodology, ranging from small observational data to a large register study. Different methodology has thus been used (Table 2). For the papers, I–III CTG recordings from a previously collected cohort were used.[33] The CTG interpretation in paper II and III were observed by Erika Gyllencreutz and Malin Holzmann.

Table 2. Methodological summary of the included papers in the thesis

	Paper I	**Paper II**	**Paper III**	**Paper IV**
Type of study	Observational study	Observational study	Observational study	Register-based study
Population or sample	Six obstetricians from two different obstetric units, 106 CTG tracings	Two obstetricians, a novel computer algorithm, 312 CTG tracings	1070 women in labor at Karolinska University Hospital undergoing FBS	127 461 low-risk deliveries, Stockholm-Gotland year 2012–2020
Exposure	Two different obstetric units with different CTG educations, national web-based CTG training vs national web-based with extended CTG training	Identification and quantification of decelerations on CTG by the two observers and the computer algorithm.	Variable deceleration duration, area and depth, average and cumulative values.	Result of admission CTG, normal or abnormal.
Outcome	Reliability in interpretation - inter- and intraobserver agreement	Agreement between computer and observers.	Lactate concentration in FBS, continuous and binary (acidemia or not, > 4.8 mmol/L cut-off for acidemia)	Proportion of SGA 10^{th} and 3^{rd} percentile. Neonatal outcomes stratified by SGA/non-SGA
Statistical analysis	Simple and weighted kappa analysis	Intraclass correlation, Spearman rank correlation and Bland-Altman analysis.	Receiver operating characteristic, area under curve. Pearson correlation.	Multiple logistic regression. Calculation of odds ratios and 95% confidence intervals.

4.1 PAPER I

Six obstetricians at two different hospitals, one university hospital and one regional hospital, were used as observers. CTG tracings (n= 106) of 60 minutes with different CTG patterns were used for interpretation, out of which 96 had been subjected to FBS due to clinically adjudged abnormal patterns. The observers documented baseline frequency (< 110 beats per minute [bpm], 110–150 bpm, > 150 bpm), variability (< 2 bpm, 2–4 bpm, 5–25 bpm, > 25 bpm), accelerations (yes or no), decelerations (absent, early, simple variable (< 60 seconds duration), late, severe variable (> 60 sec duration), combined and prolonged (> 3 min duration), contractions (< 3/10 min, 3–5/10 min, > 5/10 min), and recommended intervention (intermittent CTG, continuous CTG, FBS or urgent deliver) in a multiple-choice form attached to each case. The parameter decelerations were additionally analyzed as dichotomized into harmless (early and simple variable) and hypoxic (late, severe variable, combined and/or prolonged). The evaluation of the tracings was carried out on two occasions with at least 30 days in between.

4.1.1 Kappa analysis

The agreement between observers and within the same observer on the two occasions was analyzed by Cohen's and Fleiss' kappa. One way to measure agreement is to simply perform a proportion of agreement analysis, for example "in xx% of cases the three observers in clinic A agreed on baseline frequency". However, this method does not take chance agreement into account. When using kappa analysis, chance agreement is incorporated into the model and the kappa value (κ) is the agreement beyond chance. Furthermore, the method can weight values when ordered variables are present. For example, in terms of variability there is a greater discrepancy between reduced (2–5) and increased (> 25), than reduced and normal (5–25). The kappa statistics give a κ- value between -1 and 1, where negative values are uncommon. The value $\kappa = 0$ represents chance agreement, followed by poor agreement ($\kappa \leq 0.20$), fair agreement ($\kappa = 0.21–0.39$), moderate agreement ($\kappa = 0.40–0.59$), substantial agreement ($\kappa = 0.60–0.79$) and excellent agreement ($\kappa \geq 0.80$).

A power analysis was made to determine sample size, gaining 80% power to detect a difference between $\kappa = 0.4$ (null hypothesis) and $\kappa = 0.7$ based on previous publications in the field.[110] The two-sided significance level was set as $\alpha = 0.05$.

SAS, version 9.3 (SAS Institute Inc., Cary, NC, USA) was used for statistical analysis.

4.2 PAPER II

4.2.1 Computer algorithm

In collaboration with the Royal Institute of Technology, KTH we developed a new computerized algorithm for CTG assessment. The system can detect negative deviations from baseline, in other words decelerations, by first identifying stable regions on CTG. The baseline estimation is determined after removing non-stable regions, very short segments, and large monotonic increasing segments. Thereafter a histogram of the stable segments is analyzed and fitted with a Gaussian curve. The mean value corresponds to the baseline heart rate. Segments within two standard deviations lower and higher from baseline are included in the baseline concept. A moving filter of two-minute windows estimates the final baseline.

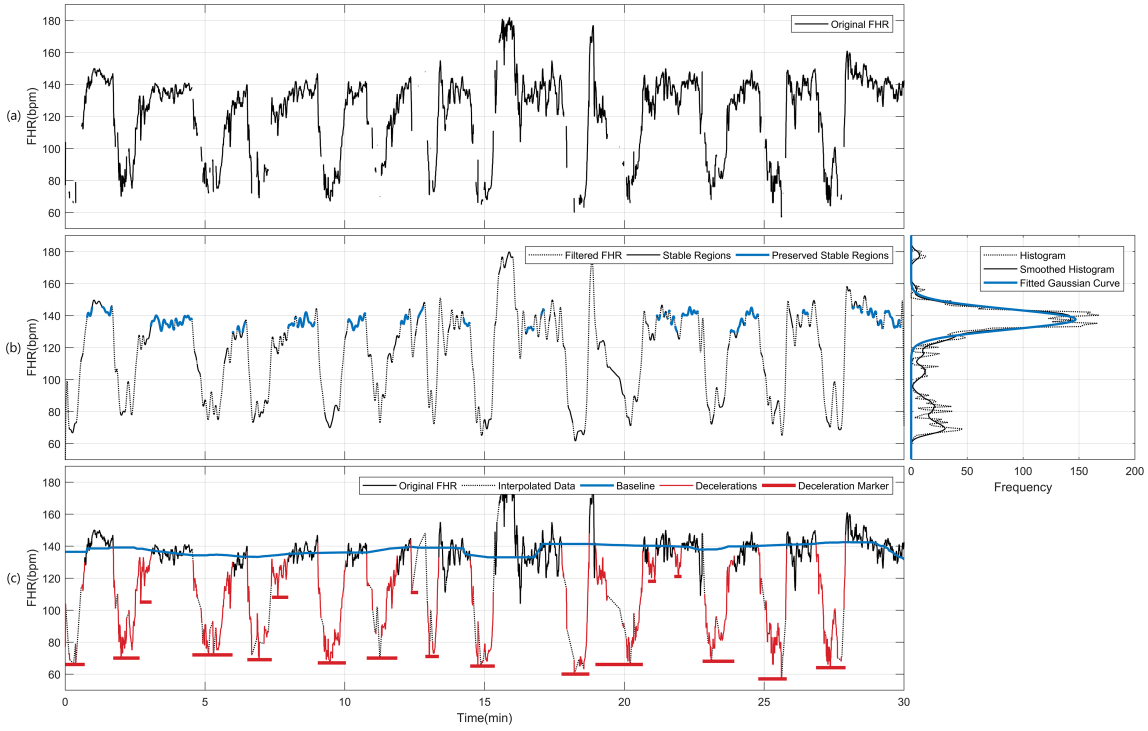


Figure 7. The process of baseline fitting and deceleration detection. (a) Original 30-min epoch of fetal heart rate (FHR). (b, left) Result of pre-processing, finding stable regions and preserved stable regions, highlighted in blue. (b, right) Preserved stable regions selected, based on fitted Gaussian distribution on the histogram of FHR. (c) Blue line, estimated baseline; red line, subsequently detected decelerations.

The algorithm detects deviations from baseline with depth > 15 bpm and width of > 15 seconds as deceleration events. Signal loss is interpolated with linear interpolation. Duration, depth and area of each individual deceleration is analyzed (Figure 7).

4.2.2 Visual observations

Two obstetricians with different levels of expertise (> 15 years and 5 years) were used for visual analysis of CTG. CTG tracings (n= 312) of 60 minutes were used for analysis, both by computer and obstetricians. Number of decelerations, total/cumulative deceleration duration (duration), area and depth were analyzed for each case. For visual interpretation, Adobe Acrobat Reader DC 2015 (Adobe Inc., San José, CA, USA) was used for detailed estimation of deceleration quantitates, by a tool in the software program (Figure 8). With its use, drawing of the individual contour of each deceleration is possible as well as measurements of duration and depth.

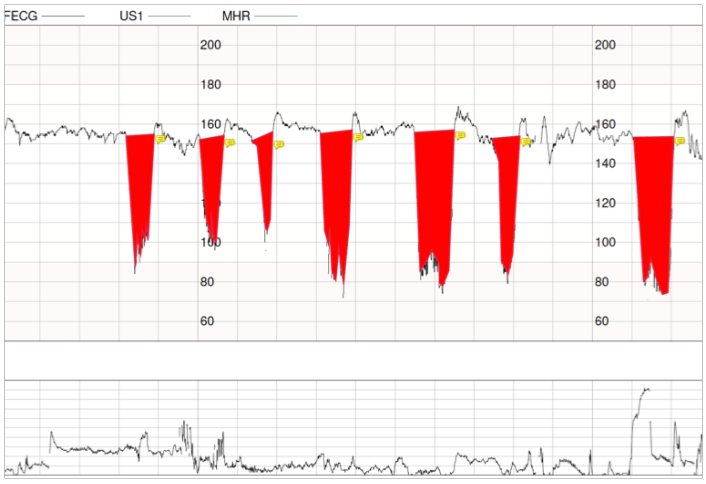


Figure 8. Estimation of deceleration area by detailed drawing of deceleration contour using Adobe Acrobat Reader

4.2.3 Statistics

STATISTICA for Windows, version 13 (Statistica, Tulsa, OK, USA) was used for all analyses.

4.2.3.1 Correlation analyses

Intraclass correlation coefficient was calculated to determine the agreement between the two observers and the computer, as well as between the two observers. Agreement was categorized as poor (0–0.5), moderate (0.5–0.75), good (0.75–0.9) and excellent (0.9–1.0) agreement. The two-way mixed effect model was used, which is preferable when there is a fixed sample of raters (not randomly selected). Additionally, Spearman rank correlation was analyzed to estimate the degree of correlation.

4.2.3.2 Bland-Altman analysis

For more thorough analysis of the computer algorithm's achievement in assessing each single deceleration, Bland-Altman analysis was made. Even though correlations can be perfect, there can be disagreement in absolute numbers, for example if one observer constantly overestimates deceleration depth. By using Bland-Altman analysis, the mean values between the two observers (computer or visual observer) for each case (here, a unique deceleration) is presented on the x-axis and the difference between the values on the y-axis, represented by a blue dot. Hence, the analysis will give the mean bias for all cases together with 95% limits of agreement (Figure 9).

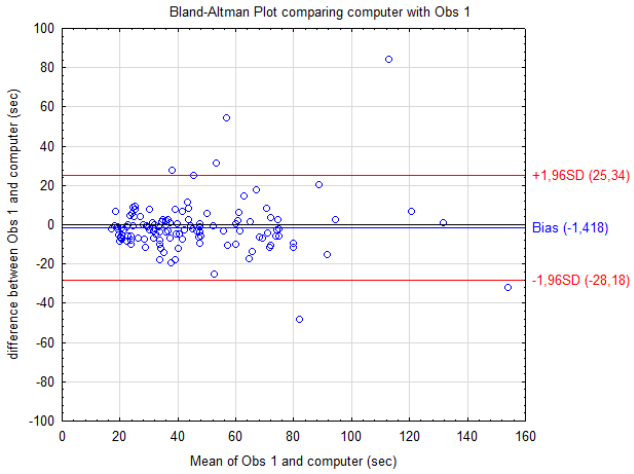


Figure 9. Example of Bland-Altman plot, each dot represents one deceleration.

Four CTG tracings were used for this analysis with approximately more than 100 unique decelerations.

4.3 PAPER III

For analysis, we used a previously collected cohort with 1070 deliveries subjected to fetal blood sampling, FBS, during labor at Karolinska University Hospital, between Feb 2009 and Feb 2011. Our focus was CTG tracings with predominantly variable decelerations (n = 507) prior to FBS. Two obstetricians, EG and MH, analyzed the CTG traces, focusing on deceleration width (duration), depth, and area for 30 and 60 minutes preceding the first (in cases of several) FBS during delivery. Adobe Acrobat Reader DC 2015 (Adobe Inc., San José, CA, USA) was used for detailed evaluation of deceleration quantitates as in paper II.

The deceleration quantitates were analyzed both as mean values per deceleration as well as cumulative values during the time period of interest (30 or 60 minutes).

Lactate concentration at FBS was used as outcome, either as a continuous value or binary with a cut-off value for acidemia above 4.8 mmol/L (earlier version of Lactate Pro™, Arkray, Kyoto, Japan).

4.3.1 Statistics

4.3.1.1 Correlation analysis

Pearson correlation was calculated between deceleration features and lactate concentration in FBS, where lactate concentration was measured as a continuous variable.

4.3.1.2 ROC AUC analysis

We determined the diagnostic value of the different deceleration quantitates to predict fetal acidemia in FBS (lactate concentration > 4.8 mmol/L) by using ROC AUC analysis. The ROC curve includes sensitivity on the Y-axis and 1-specificity on the X-axis (Figure 10). The area under the curve, AUC, represents the strength of the diagnostic test. The perfect diagnostic test for a binary outcome will have 100% sensitivity and 100% specificity, hence the AUC will be 1.0. An AUC representing a diagnostic performance of chance, e.g., flipping a coin will have an AUC of 0.5.

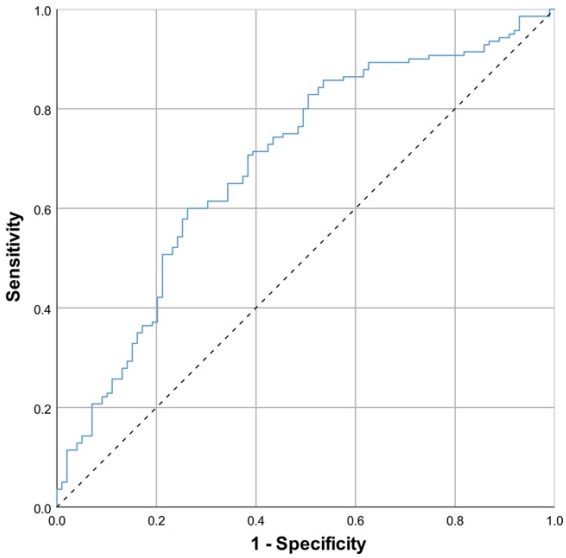


Figure 10. Example of ROC AUC, dashed line represents AUC 0.5, blue line AUC of 0.70.[111]

4.4 PAPER IV

The last study is a register-based study from the Stockholm-Gotland Perinatal cohort 2012–2020 including 192 314 women. Inclusion criteria were singleton gestation, vertex lie, gestational age > 34 weeks. Exclusion criteria were elective cesarean delivery, placenta previa, CNS anomaly, fetal arrhythmia, induction of labor (other than due to suspected fetal hypoxia). The cohort was additionally subdivided into low-risk and high-risk depending on predefined maternal and obstetric conditions such as maternal morbidity, detected fetal growth restriction, oligohydramniosis, macrosomia, infection, placenta abruptio, preterm and postterm gestational age. Information on CTG admission results was extracted from the register, classified as normal, abnormal, not performed, and missing.

The outcome of interest was the proportion of neonates born small for gestational age 3rd and 10th percentile based on the result of admCTG in the low-risk population. Secondary outcome was composite severe and moderate adverse neonatal outcome for those neonates born within six hours after admission, stratified by SGA and AGA/LGA, and result of admCTG (normal or abnormal). Composite severe adverse neonatal outcome included at least one of following: Apgar scores < 4 at 5 minutes, HIE grade 2–3, neonatal seizures, neonatal, and intrapartum death. Composite moderate adverse neonatal outcome was defined as: Apgar scores < 7 at 5 minutes, umbilical artery pH < 7.00, metabolic acidosis defined as umbilical artery pH < 7.05 and BD ≥ 12 mmol/L, or the need of advanced neonatal resuscitation.

Odds ratios (OR) with 95% confidence intervals (CI) were calculated for the main outcome. Secondary outcomes were analyzed by calculating adjusted OR (aOR) with multiple logistic regression. This model included the following potential confounders; parity (dichotomized as nulliparous and parous), BMI (continuously) and calendar year (dichotomized as before or after year end 2016, when a new CTG classification system was launched). IBM SPSS Statistics, version 28, was used for the statistical analysis.

4.5 ETHICAL CONSIDERATION

Ethical approval was obtained for all studies in the thesis. In the original cohort, including 1070 women at study start, consent was obtained from all patients in the study. However, the research group recognized that it would be difficult to reach enough power with only the included women. Hence, an additional ethical approval was sought and approved, where all women subjected to FBS during the study period were included, with assumed consent to participate. Thus, medical records from all women could be analyzed. One can discuss whether this was ethically correct since there had been an ambition to obtain informed consent, and it was not possible to fulfil this. However, if the study had continued without the supplementary ethical approval, the insufficient power in an already started study would introduce another ethical dilemma.

In studies I-III performed on the above cohort, data are presented at a group level and none of the participating individuals are possible to identify. During the process of the studies there has been careful precautions with handling of the data.

In the last paper IV, we have included laboring women in the Stockholm-Gotland region during 2012–2020 through a register. When presenting the data, which includes very rare outcomes, the results are aggregated at a group level to decrease the risk of identification and outcomes involving < 5 individuals per group have not been presented with specific numbers and percentages.

When I consider different aspects of the research field and my projects in particular, I argue that the potential benefits of the work outweigh the potential risks. If we were to gain new knowledge in the field of CTG including new computerized assessments, we could make the

tool more effective in finding the fetuses at risk of acidemia during labor. This could in fact benefit the study participant in her future pregnancies and other women and their fetuses.

RESULTS

4.6 PAPER I

4.6.1 Interobserver agreement

The interobserver agreement was moderate to substantial at both departments, with κ from 0.41 to 0.76. The department with extended education performed significantly better in terms of variability and accelerations (Table 3). The reference department achieved better agreement for contractions. Overall, at both departments, agreement was lower for variability and decelerations than baseline frequency, accelerations, and contractions. Decelerations dichotomized as harmless or hypoxic reached better agreement compared to current categorization.

Table 3. Level of inter-observer agreement, kappa values first and second assessment

CTG parameter	Department with regular education (95% CI)	Department with extended education (95% CI)	p-value
Baseline frequency	0.71 (0.63–0.78)	0.76 (0.69–0.83)	0.34
Variability	0.41 (0.35–0.48)	0.53 (0.46–0.60)	0.01
Accelerations	0.60 (0.52–0.68)	0.72 (0.65–0.80)	0.03
Decelerations	0.44 (0.41–0.48)	0.42 (0.39–0.46)	0.50
Dec dichotomized*	0.63 (0.55–0.71)	0.63 (0.55–0.71)	0.96
Contractions	0.69 (0.63–0.74)	0.58 (0.52–0.63)	< 0.01

*=early and simple variable decelerations versus late, severe variable, combined and prolonged decelerations

4.6.2 Intraobserver agreement

Reliability in terms of intraobserver agreement was substantial to excellent at both departments (Table 4). The department with extended education achieved significantly better results for baseline frequency, variability, and accelerations. When decelerations were dichotomized the level of agreement was increased.

Table 4. Level of intraobserver agreement, kappa values

CTG parameter	Department with regular education (95% CI)	Department with extended education (95% CI)	p-value
Baseline frequency	0.78 (0.72–0.85)	0.93 (0.90–0.97)	<0.01
Variability	0.66 (0.58–0.74)	0.80 (0.73–0.86)	<0.01
Accelerations	0.72 (0.64–0.79)	0.92 (0.87–0.96)	<0.01
Decelerations	0.65 (0.60–0.71)	0.65 (0.59–0.70)	0.86
Dec dichotomized*	0.78 (0.71–0.85)	0.76 (0.69–0.83)	0.71
Contractions	0.74 (0.65–0.84)	0.81 (0.74–0.88)	0.25

*=early and simple variable decelerations versus late, severe variable, combined and prolonged decelerations

4.7 PAPER II

4.7.1 Analysis of 312 cases

In the full cohort of 312 cases, intraclass correlation was excellent (0.82-0.91) and Spearman rank correlation strong (0.82-0.91) between the performance of the observers and of the computer algorithm regarding detection and quantitation of decelerations in terms of number of decelerations, as well as cumulative deceleration duration, depth and area for each case.

4.7.2 Analysis of four cases with more than 100 decelerations

In the more detailed analysis of four cases with > 100 decelerations in total, the analysis of median values for deceleration measures (duration, depth and area) with Kruskal-Wallis test showed no difference among the three observers (Obs 1, 2 and computer) in terms of deceleration duration and area. However, the assessment for deceleration depth was significantly different between the observers with a p-value of 0.02 (Table 5).

The Bland-Altman analysis, comparing the individual observers with each other and with the computer, reported low mean bias for all parameters: duration, depth, and area. The 1.96 standard deviation (SD) was comparable between computer and observers and between observers. For example, for deceleration duration 1.96 SD ranged from -28.2 to 25.3 (Obs 1 vs computer), -30.3–27.6 (Obs 2 vs computer) and -25.3–24.7 (Obs 2 vs Obs 1).

Intraclass correlation in the detailed analysis was excellent between the computer and the two observers as well as between the observers. (Table 5).

Table 5. Characteristics and Bland-Altman analysis of deceleration measures (four cases)

	Duration (sec)		Depth (beats/min)		Area (beats)	
Number and characteristics of decelerations[a]						
Observer 1 (n=127)	39 (30–59)		40 (26–53)		14 (8–27)	
Observer 2 (n=128)	40 (29–60)		44 (38–61)		16 (9–30)	
Computer (n=138)	35 (25–55)		46 (31–60)		13 (7–26)	
p-value[b]	0.13		0.02		0.14	
Bland-Altman analysis of deceleration measures	**Bias**	**±1.96 SD**	**Bias**	**±1.96 SD**	**Bias**	**±1.96 SD**
Observer 1 vs Computer	-1.4	-28.2–25.3	5.1	-9.1–19.2	0.1	-11.4–11.6
Observer 2 vs Computer	-1.4	-30.3–27.6	0.7	-10.8–12.3	-1.7	-11.8–8.5
Observer 2 vs Observer 1	-0.3	-25.3–24.7	-4.4	-16.7–7.9	-1.7	-13.9–10.5

[a]Median (25th–75th percentile), [b]Kruskal-Wallis test

4.8 PAPER III

There were more frequent decelerations, a larger cumulative deceleration area, duration and depth when comparing the group of fetuses with acidemia (lactate concentration > 4.8 mmol/L) at FBS to those without acidemia. For mean deceleration depth/deceleration there was no significant difference between the groups, but a larger mean area and duration per deceleration in the group of fetuses with acidemia compared to no acidemia.

Pearson correlation between the different deceleration quantitates and lactate concentration at FBS were weak to moderate. The highest correlations were achieved by cumulative deceleration area and duration, r = 0.33 and 0.35 respectively in contrast to cumulative deceleration depth, r = 0.21.

The predictive value for the different deceleration quantitates and acidemia at FBS (lactate concentration > 4.8 mmol/L) was carried out by ROC AUC (Table 6). The highest AUC was obtained for cumulative deceleration area and duration 60 minutes preceding FBS. Mean depth/deceleration was no better than chance in predicting fetal acidemia.

Table 6. Variable deceleration features and prediction of intrapartum acidemia in FBS

	AUC 30 min (95% CI)	**AUC 60 min (95% CI)**
Total dec area, beats	0.671 (0.588-0.754)	0.682 (0.595-0.769)
Total dec duration, sec	0.678 (0.593-0.762)	0.683 (0.595-0.771)
Total dec depth, bpm	0.632 (0.550-0.715)	0.631 (0.546-0.716)
Area/deceleration	0.621 (0.540-0.702)	0.627 (0.549-0.706)
Duration/deceleration	0.622 (0.542-0.702)	0.630 (0.556-0.704)
Depth/deceleration	0.558 (0.479-0.637)	0.559 (0.475-0.642)
Number of decelerations	0.624 (0.542-0.705)	0.620 (0.534-0.707)

Intrapartum acidemia defined as lactate concentration > 4.8 mmol/L at FBS
AUC, area under receiver operating characteristic curve; *bpm*, beats per minute; *CI,* confidence interval; *dec,* deceleration; *sec,* seconds; *min,* minutes

With a cut-off value for cumulative deceleration area of 250 beats, 71% of the acidemic fetuses had cumulative deceleration area above this threshold compared to 43% of the fetuses without acidemia, OR 3.2 (95% CI 1.7–6.1). Hence, the sensitivity was 0.71 and specificity 0.57.

4.9 PAPER IV

From the Stockholm-Gotland Perinatal cohort the eligible study population consisted of 153 061 women after inclusion and exclusion criteria were fulfilled. Of these 25 555 pregnancies were considered high-risk and 127 461 as low-risk. Of those women with information on admCTG result (n = 110 681), 4.9% were classified with abnormal admCTG, 93.9% with normal admCTG and in 1.2% admCTG was not performed.

4.9.1 Proportion of SGA and result of admCTG

In the low-risk population, SGA 10th and 3rd percentile was more common in the abnormal admCTG group compared to the normal, 18.6% vs 9.7%, OR 2.1 (95% CI 1.9–2.3) and 6.5% vs 2.2%, OR 3.1 (95% CI 2.7–3.5), respectively.

4.9.2 Neonatal outcomes in SGA fetuses with abnormal admCTG

Neonatal outcomes were analyzed for those neonates born within six hours after admission in the low-risk population. The risk of composite severe adverse neonatal outcome was 20-fold higher in the group of fetuses presenting with abnormal admCTG and born SGA 10th percentile (1.2%) compared to those with normal admCTG and AGA or LGA (0.05%), aOR 23.7 (95% CI 9.8–57.3). Corresponding aOR for composite moderate adverse neonatal outcome was aOR 5.8 (95% CI 4.0–8.4) affecting 6.6% of neonates with abnormal admCTG/SGA vs 1.0% of neonates with normal admCTG/AGA or LGA. Operative delivery due to fetal distress, cesarean or vacuum, was more common in the abnormal admCTG/SGA group compared to the normal admCTG/AGA/LGA; 23.2% compared to 1.3%, aOR 15.6 (95% CI 12.5–19.6) (Table 7 and Figure 11).

Table 7 Fetal and maternal outcome among women assessed as low-risk and delivered within six hours from admission, crude odds ratio (OR), n =54432

	Normal admCTG (n = 51495) †		**Abnormal admCTG (n = 2937) ‡**	
	AGA/LGA	SGA	AGA/LGA	SGA
	n = 47004	n = 4463	n = 2370	n = 564
	n (%) OR (95% CI)	n (%) OR (95% CI)	n (%) OR (95% CI)	n (%) OR (95% CI)
Composite severe adverse neonatal outcome*	21 (0.05%) 1.0 (ref)	6 (0.1%) 3.0 (1.2–7.5)	10 (0.4%) 9.5 (4.5–20.2)	7 (1.2%) 28.2 (11.9–66.5)
Apgar <4 at 5 minutes	8 (0.02%) n.a.	<5 (0.1%) n.a.	7 (0.3%) n.a.	<5 (0.9%) n.a.
Neonatal seizures	12 (0.03%)	<5 (0.1%) n.a.	<5 (0.2%) n.a.	<5 (0.9%) n.a.
HIE grade 2-3	<5 (0.01%)	<5 (0.1%) n.a.	<5 (0.2%) n.a.	<5 (0.9%) n.a.
Neonatal death	<5 (0.01%)	<5 (0.1%) n.a.	<5 (0.2%) n.a.	<5 (0.9%) n.a.
Intrapartum death	<5 (0.1%)	<5 (0.1%) n.a.	<5 (0.2%) n.a.	<5 (0.9%) n.a.
Composite moderate adverse neonatal outcome*	447 (1.0%) 1.0 (ref)	59 (1.3%) 1.4 (1.1–1.9)	88 (3.7%) 3.6 (2.8–4.5)	37 (6.6%) 6.5 (4.6–9.3)
Apgar <7 at 5 minutes	72 (0.2%) 1.0	13 (0.3%) 1.9 (1.1–3.4)	27 (1.1%) 7.5 (4.8–11.7)	11 (2.0%) 13.0 (6.9–24.7)
pH <7.00	92 (0.4%) 1.0	18 (0.7%) 2.1 (1.3–3.5)	31 (2.1%) 6.0 (4.0–9.1)	8 (2.2%) 6.5 (3.1–13.5)
pH <7.05 and BD ≥ 12	216 (0.8%) 1.0	28 (1.1%) 1.4 (0.9–2.1)	41 (2.7%) 3.4 (2.4–4.8)	9 (2.5%) 3.1 (1.6–6.1)
Advanced neonatal resuscitation**	222 (0.5%) 1.0 (ref)	28 (0.6%) 1.3 (0.9–2.0)	58 (2.5%) 5.3 (3.9–7.1)	30 (5.3%) 11.8 (8.0–17.5)
Composite obstetric outcome*	606 (1.3%) 1.0 (ref)	136 (3.1%) 2.4 (2.0–2.9)	240 (10.1%) 8.6 (7.4–10.1)	131 (23.2%) 23.2 (18.8–28.6)
Cesarean delivery due to fetal distress	96 (0.2%) 1.0	36 (0.8%) 4.0 (2.7–5.8)	137 (5.8%) 30.3 (23.0–39.0)	89 (15.8%) 91.6 (67.7–123.9)
Instr delivery due to fetal distress	517 (1.1%) 1.0	103 (2.3%) 2.1 (1.7–2.6)	104 (4.4%) 4.1 (3.3–5.1)	43 (7.6%) 7.4 (5.4–10.2)

* = at least one of the outcomes below

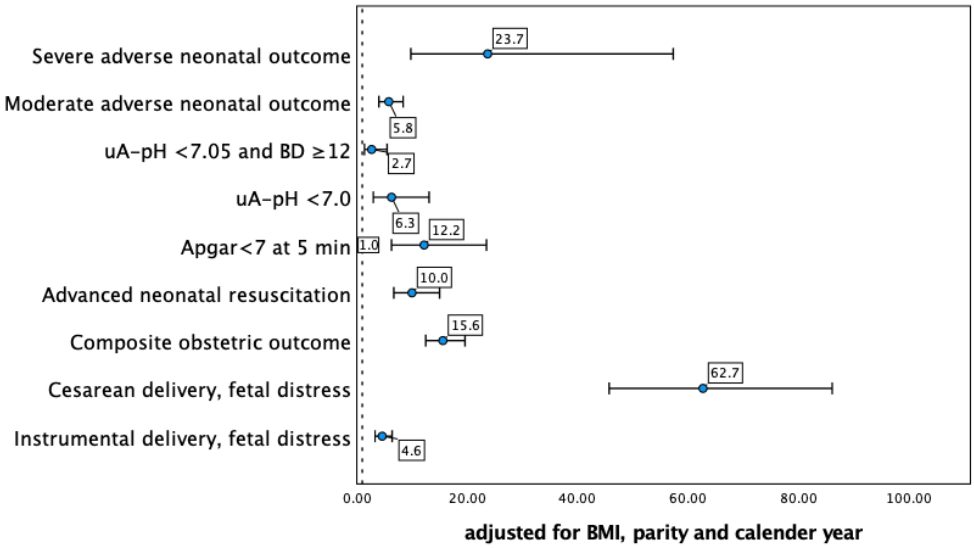
** correction of acidosis, adrenaline use, CPAP, extra oxygen supplement, heart compressions

BD, base deficit in mmol/L; HIE, hypoxic ischemic encephalopathy; instr, instrumental; OR, odds ratio

† 28 neonates and ‡ 3 neonates missing information on birth weight

For outcome <5 ORs are not calculated, and exact numbers are not given

Figure 11. Adjusted OR comparing abnormal admCTG/SGA (10th percentile) vs normal admCTG/AGA or LGA, neonates delivered < 6 hours from admission, low-risk deliveries, n = 47 568



5 DISCUSSION

5.1 MAIN FINDINGS

When comparing two departments with two different rationales for CTG training, extended education might result in better inter- and intra-observer agreement compared to the national routine web-based education. Nevertheless, CTG reliability was better than expected at both departments reaching moderate to substantial agreement in terms of inter-observer agreement and substantial to excellent agreement for intra-observer reliability. When decelerations were dichotomized into harmless and hypoxic, reliability increased compared to current categorization of decelerations (early, simple variable, late, severe variable, combined, prolonged).

A new computerized algorithm has been invented in collaboration with engineers at the Royal Institute of Technology, KTH. The computerized algorithm can accurately detect decelerations and measure deceleration duration, area, and depth. When comparing intraclass correlations between computer and two individual observers the correlations were excellent (0.90-0.98) and comparable to that between the two observers (0.93–0.97). When estimating bias between computer and observers using Bland-Altman analysis the bias was low and comparable to that between the two observers.

Scrutinizing the predictive ability of different deceleration features and acidemia during labor, measured as lactate concentration at FBS, renders that deceleration area and duration are better predictors of intrapartum acidemia compared to deceleration depth. When introducing a cutoff-value for cumulative deceleration area of 250 beats per 30 minutes, breaching this threshold was associated with three-fold higher odds of intrapartum acidemia compared to < 250 beats, OR 3.20 (95% CI 1.67–6.13).

In a low-risk population including > 100 000 women, the proportion of undetected SGA, 10th percentile is approximately 10%. When presenting with abnormal CTG at admission the risk of SGA is twice as high compared to when a normal tracing is achieved. Moreover, the group of neonates born SGA with abnormal admCTG is at a considerably higher risk of severe and moderate adverse neonatal outcomes compared to the group of AGA/LGA neonates with normal admCTG.

5.2 INTERPRETATION

CTG is engaged with several shortcomings: low reproducibility[70-72,112], and low specificity[75,76] with a high proportion of false positive CTG tracings indicating fetal acidemia. In addition, it has been difficult to prove the neonatal benefit of CTG, both when used during labor [64] and as admCTG.[84-86,113,114]

However, we found an unexpectedly high reliability in CTG interpretation at two different departments with two different strategies for CTG education. In our study we have focused on individual parameters in CTG interpretation: baseline frequency, variability, accelerations, decelerations, and contractions. The observers were not asked to classify CTG as normal, suspicious, pathological or pre-terminal as opposed to other studies evaluating CTG reliability.[70,72,112] Nonetheless, one more study has demonstrated good interobserver reliability both for individual CTG parameters as well as classification according to NICHD.[74]. In our material, all CTG tracings were collected from parturients subjected to FBS, extracting the CTG tracing immediately preceding FBS. Hence, these tracings are considered at the least suspicious. Previous studies have observed that CTG reliability is lower for intermediary tracings compared to normal and very pathological tracings.[71,112,115]

The parameters with the lowest CTG agreement were variability and decelerations which is in line with previous findings.[73,116] When dichotomizing decelerations into harmless and hypoxic agreement increased compared to the current sub-classification of decelerations. Ayres-de-Campos and Bernardes showed poor interobserver agreement distinguishing between early, late and variable decelerations.[117] They concluded that it might be of greater importance to evaluate width and depth of decelerations than to classify them into subgroups.

To overcome the obstacle with the arguably poor interobserver reliability, computerized assessment of CTG has been proposed as an attractive alternative. Several systems have been developed.[118] But so far, the clinical use of computer driven CTG interpretation has been scarce.[14,91] In addition, a computerized method is necessary to enable continuous assessment of the deceleration area. We developed a new computerized algorithm to detect and measure decelerations with focus on variable decelerations. A low bias and a high degree of correlations were found between computer and observers and between the two observers.

Previous studies assessing computer-calculated deceleration area have reported an association between deceleration area and neonatal acidemia.[55-57] Other studies have used the calculation deceleration width × depth / 2 as an estimation for deceleration area.[34,54,119,120] This estimation most likely underestimates the accurate deceleration area, since decelerations rarely are triangular in shape. Nonetheless, deceleration area was the best predictor of neonatal acidemia in the studies.[54,57,119]

We argue that deceleration area and duration are better predictors of intrapartum acidemia than deceleration depth. However, previous research suggest that the deceleration depth (also referred to as amplitude), reflects the severity of fetal hypoxia.[121] We suggest that it is favorable to assess cumulative deceleration quantitates (area and duration) during 30 or 60 min rather than mean values per deceleration, since the former takes the number of decelerations into account and estimates the summed burden of beat loss. However, the ROC AUC for cumulative deceleration area in our material was quite modest (0.671–0.682) compared to previous studies 0.702–0.83.[54,57,119]

Previous studies have used a cutoff of 250 beats and studied association with neonatal acidemia.[54,57,119] When we used the same cutoff, we found an increased risk of intrapartum acidemia. When passing this threshold, 71% of fetuses with acidemia (lactate concentration > 4.8 mmol/L) had a beat loss > 250 beats compared to 43% of fetuses without acidemia.

Further, the previous studies have not distinguished between early, variable, and late decelerations. We focused on variable decelerations which is the most frequent deceleration during labor,[122] with the knowledge that even shallow late decelerations can be correlated to fetal acidemia.[123] The magnitude of variable decelerations might thus be of greater importance in predicting fetal acidemia compared to the magnitude of late decelerations. However, new evidence suggests that during labor it is of less importance to distinguish between variable and late decelerations. The late decelerations arising during labor might be of another nature than late decelerations present in the earliest phase of labor or latency phase.[41,121] A previous study from our group observed a similar incidence of acidemia at FBS among fetuses presenting with severe variable decelerations as when presenting with late decelerations.[33] In the study by Cahill et al, the ROC AUC for repetitive late and variable decelerations predicting neonatal acidemia were comparable, 0.78 and 0.79 respectively.[34] For future studies, the computer algorithm in our study has to include the temporal relationship to the uterine contractions and/or time from start of contraction to nadir, to distinguish late from variable decelerations. Alternatively, all decelerations should be analyzed together, as in previous studies.

In terms of admCTG some argue against its use and some national guidelines, for example in the UK and Norway do not recommend it in low-risk deliveries.[48,79-81] However, this strategy has been debated.[124-126] The main objection is that the RCTs, which the guidelines are based on, are of insufficient sample size to detect differences in rare events such as HIE. Consequently, some guidelines, for example Australia and New Zealand, leave it to the attending clinician or institution to decide whether to use admCTG or not, individualized for the patient in question.[127] Nevertheless, contrary to guidelines, admCTG is performed in the majority of low-risk deliveries in the UK and Norway.[85,128] In Sweden, the recommendation is admCTG for all women.[129] One of the objections against admCTG is that it increases the risk of continuous CTG compared to IA.[85] However, in the RCTs performed the proportion of abnormal admCTG was high, 22–32%[84,114] in contrast to our study with 4.9% abnormal CTG in the low-risk group. Hence, this could explain the increased risk of continuous CTG in their material. Moreover, in the largest RCT, early amniotomy was routine procedure before randomization, establishing information on meconium or not and in case of meconium the woman was not considered low-risk and excluded from analysis.[84] In Sweden, early amniotomy is not a current routine, and therefore this information on one of the fetal risk factors, meconium or not, is not present at the time of admission.

We observed a higher risk of neonatal complications among SGA neonates. This is in line with previous findings.[130,131] In addition, the combination of abnormal admCTG and SGA increased the risk of neonatal complications even more. In a study by Akhavan et al, women

with abnormal admCTG were compared to a group with normal admCTG. The proportion of fetal growth restriction was 17% when admCTG was abnormal and 3% when normal. The neonates with abnormal admCTG were also subjected to lower pH, lower Apgar scores, increased risk of cesarean delivery, and admission to neonatal intensive care unit.[132] Parts et al studied > 40 000 deliveries and found that admCTG was beneficial for the decision to intervene in 74% of women who underwent cesarean delivery within one hour after admission due to suspected fetal distress. When assessing the CTG patterns, the presumed ability to find the ominous or abnormal patterns by auscultation alone was 28%. In the study group 35% was SGA 10th percentile compared to 11.1% in the control group of women not delivered by cesarean section within one hour after admission. Lovers et al studied admCTG among > 27 000 parturients and analyzed three groups; neonates delivered with cesarean delivery due to suspected fetal distress within two hours after admission, between 2–5 hours after admission, and the rest of deliveries.[133] As in our study, they aimed to address the original idea with admCTG – to detect fetuses who are already compromised or vulnerable at the time of admission.[78] They studied neonates born with severe compromise compared to uncompromised neonates and found that the risk of SGA was approximately three times higher in the severely compromised fetuses regardless of time from admission to delivery. Tachycardia > 150 bpm, reduced long-term and short-term variability, prolonged decelerations and decelerative capacity was more common among the compromised neonates than the non-compromised.

Which admCTG patterns that are possible to identify by auscultation alone is another question. The NICE guidelines recommend auscultation at least one minute immediately after one contraction and every 15 minutes thereafter.[48] Hence, there is a risk of missing prolonged non repetitive decelerations. Moreover, there are difficulties to identify non-reactive fetal heart rate patterns with decreased variability, and discreet decelerations. Decelerations need to be of great amplitude to be auscultated. For example, see figure 6 above. Also, detecting late decelerations might be difficult when auscultating only one minute after one contraction; there are case reports of misinterpretation of baseline during late decelerations with neonatal encephalopathy as a result.[134] This has been pointed out by Sholapurkar, who recommends auscultation to begin before the end of one contraction and continuing to the beginning of the next coming contraction.[135]

5.3 STRENGTHS AND LIMITATIONS

To our knowledge we have performed the first study evaluating the effect of two different education strategies on CTG interpretation, with focus on inter- and intra-observer reliability. Nevertheless, there were differences between the two obstetric units besides educational strategies – for example, one unit is a university hospital and the other a secondary regional hospital with fewer deliveries per year. Another strength is that we have used tracings that are presumably more difficult to interpret, the intermediary tracings. Still, relatively good inter- and intraobserver reliability was observed. We did not ask for classification of the CTG tracings which earlier interobserver studies have included. This could in some degree explain

our satisfactory agreement – it might be more difficult to agree on classification (normal, suspicious, pathological) than the individual parameters of CTG. Nonetheless, there are a wide range of classification systems internationally and results from interobserver studies might be difficult to generalize to other systems.

In addition, we have not seen previous studies investigating the precision of a computerized algorithm in terms of identifying and quantitating decelerations in detail as in our study. As discussed above, it is necessary to achieve an adequate computerized tool if deceleration area is to be measured during labor in real-time. We have not used CTG tracings immediately prior to delivery, which might be more difficult to assess, both for computer and observers; the baseline could be harder to distinguish, and hence the decelerations more difficult to measure. Also, signal loss is more common at the end of the delivery. Thus, studies assessing the performance of the computerized algorithm in the penultimate phase of the delivery is important. For the Bland-Altman analysis we used only four CTG tracings. However, since we focused on variable decelerations which by definition differ from one another in the same fetus, and since each individual deceleration was regarded as one case, we achieved more than 100 individual decelerations for analysis. Nevertheless, we performed correlation analysis between computer and observer in an additional 312 CTG tracings with more than adequate agreement. Still, assessments between the two observers were not feasible. In other studies where a computerized system has been validated, the observers have been senior experts[136,137] as opposed to our study consisting of one senior and one junior obstetrician. Anyhow, the bias between the two observers was low and correlations high.

In the third study we used a large material consisting of 507 CTG tracings. No previous study has investigated the effect of deceleration area on intrapartum acidemia, here measured as lactate concentration at FBS. All previous studies have focused on neonatal acidemia in the newborn. We reflect that assessing the CTG tracing immediately prior to delivery is associated with some disadvantages. Often there can be a lag time between last CTG registration and birth, due to cesarean delivery or CTG signal loss, during the active second stage. The use of betamimetics for intrauterine resuscitation before cesarean delivery can also alter acid-base condition in relation to CTG pattern.[138,139] Therefore, we used lactate concentration during labor as outcome. Hence, the CTG tracing analyzed is directly adjacent to the time of FBS. Since our cohort consists of fetuses who have undergone FBS, we propose that it constitutes a high-risk group for acidemia. The results are thus not generalizable to an unselected population. We suggest that this fact could explain the lower AUC for acidemia in our study compared to other studies – the difference in deceleration area between the acidemic vs non-acidemic group might be larger in previous studies, generating higher AUCs.

The register-based study population in paper IV is large, including 127 461 low-risk women. It is not possible to gain this number of participants in an RCT, which means that we were able to scrutinize rare adverse complications as a composite outcome and demonstrate significant differences. We have also assessed outcomes in women undergoing routine labor

care, using admCTG as intended before anamnesis and examination, in contrast to the previous RCTs with early amniotomy and very high proportions of abnormal admCTG compared to ours. We have also focused on the original intention with admCTG, to detect vulnerable fetuses with ongoing risk of compromise at admission, why we have restricted our analysis to fetal outcomes occurring during the first six hours after admission and not later as earlier studies have. We have not been able to rule out which neonates who suffered from true fetal growth restriction and therefore we used SGA 10th percentile as a proxy. Indeed, it could be suggested that using 3rd percentile would be more accurate. However, internationally the 10th percentile has been more common as a definition for SGA. Nevertheless, we could demonstrate large differences in neonatal outcomes when using this cutoff. There is missing data concerning the result of admCTG (13%) and umbilical cord gas analysis (45%) in the low-risk deliveries, which introduces potential bias. However, we argue that in cases of compromised neonates, the umbilical cord gas collection is prioritized. The missing cord gases might thus mainly concern vigorous neonates. We also performed analyses of the neonates with missing information on admCTG and found that they had substantially lower adverse neonatal risks than SGA neonates with abnormal admCTG.

6 CONCLUSIONS

Inter- and intra-observer agreement in CTG interpretation performed at two different obstetric units was better than expected, compared to previous studies. An extended CTG education combining lectures, web-based training, and oral examinations, may lead to better reliability in interpretation.

Computerized assessment of CTG, by a newly developed algorithm, is as accurate as visual assessment of CTG in terms of adequately detecting decelerations and measuring deceleration duration, depth, and area adequately.

Cumulative deceleration area and duration are better predictors of intrapartum fetal acidemia compared to cumulative deceleration depth. When using a cutoff of > 250 beats for 30 minutes cumulative deceleration area, the odds of fetal acidemia was three-fold higher compared to < 250 beats.

Among presumed low-risk pregnancies there is a substantial proportion of undetected SGA fetuses who more often present with abnormal CTG at admission compared to non-SGA fetuses. These SGA fetuses also experience higher risks of severe and moderate adverse neonatal outcomes, especially when admCTG is abnormal.

7 POINTS OF PERSPECTIVE

After more than half a century of its use, there is still much knowledge to gain in CTG and its role to predict outcome in the newborn, ultimately enabling interventions before irreversible hypoxic complications arise.

The computerized algorithm could be refined, and further developed. Future studies should include a larger material and exploring outcomes in the newborn, which is the ultimate goal with CTG, to reduce negative outcomes in the neonate.

It would be meaningful to elucidate the difference between the variable and the late decelerations that emerge during labor, both in terms of pathophysiology but also predictive capacity regarding fetal acidemia. Does the association between deceleration area and acidemia differ between the two types of decelerations?

Another aim for future studies is to assess the effect of baseline increment in combination with deceleration area and fetal/neonatal acidemia, i.e., to assess the deviation from the baseline at the beginning of labor in addition to cumulative deceleration area with acidemia.

With respect to admCTG, assessment and interpretation of abnormal CTG tracings among SGA neonates delivered close to admission could be analyzed. Focus would be to assess whether the abnormal admCTG could have been detected with auscultation alone.

8 ACKNOWLEDGEMENTS

I have numerous people to thank who have helped me on the way and made this thesis possible.

My dearest supervisor, **Malin Holzmann**. You have been invaluable as a guide and companion, leading me in the right direction, nurturing me from the start, supporting me on the way and eventually letting me try my own wings. This long-distance PhD education would not have been possible without your tireless work. Also, thank you for coping with my last-minute nature and not fully organized structure as a person.

My co-supervisor **Pelle Lindqvist**, with impressive capacity, and scientific skills. Thank you for insightful comments and I love your preference for spontaneous phone calls when you have come up with an idea no matter time, place, or day. You have also introduced me to the genius use of blue print in science.

My co-supervisor, **Ingela Hulthén Varli**, the clinical professional, who always has something new in mind and undefeated in finding a new angle when revising manuscripts. You have most kindly answered my clinical questions when I am lost. It has been a pleasure to get to know you.

My co-authors **Ke Lu, Kaj Lindecrantz, Farhad Abtahi** the skilled engineers from KTH, **Lennart Nordström,** The Obstetrician and **Kari Johansson**, KEP, for professional and fun collaboration.

My mentor **Thomas Drevhammar**, backing me up and coming with street-smart advice when the wind has been blowing.

Anna Lindam, FoU statistician for valuable comments on the methods used in the thesis.

Iréne Hoglert and **Kristina Nordquist**, heads of Kvinnosjukvården, Östersund for making it possible to conduct research beside the clinical work. You have always been positive and supportive of my research.

My dear colleagues, old and new, at Kvinnosjukvården, Östersund. I have great fun with you and you're my friends and extra family. Thank you, **Eva Spetz**, for giving me the opportunity to leave the screws and broken bones and try a new path. I do not regret it. You taught me how to become an obstetrician.

All midwives and assistant nurses at work. Some of you have seen and helped me grow up. All of you make my days developing and fun.

FoU Jämtland Härjedalen for generous financial support from the start. Thank you for bringing the academic possibility to Östersund Hospital.

My research friend **Viktor Grut** for inspiring talks about science and for your review of the thesis.

My dear friend and gifted artist/surgeon **Elisabeth Abdon**, who illustrated the cover of this thesis.

My colleague and friend **Anna Persson** for geniune encouragement in both work and life. I can always trust on you for support.

My longtime friend **Charlotta Larsson** who shares my interest in research and more importantly a million other things. You supported me during this arduous process of finishing my thesis. Your advice regarding small and big things are invaluable.

My professional and private eternal companion **Ann Edblad Svensson**, following me on the way wherever I go.

My wonderful **friends,** for sharing good and on occasion bad times. I am so grateful for having you in my life!

My parents **Joachim** and **Tuula** and my brother **Henrik** for great support and help during the years.

My wonderful boys, **Vilhelm**, **August** and **Viktor**. You are my everything.

My love **Christian**, always there during sun, rain and thunder, like a pillar. Thank you for your endless support and always believing in me when I am in doubt.

9 REFERENCES

1. Salhany JM, Mizukami H, Eliot RS. The deoxygenation kinetic properties of human fetal hemoglobin: effect of 2,3-diphosphoglycerate. *Biochem Biophys Res Commun* 1971; **45**(5): 1350-6.
2. Nicolaides KH, Economides DL, Soothill PW. Blood gases, pH, and lactate in appropriate- and small-for-gestational-age fetuses. *Am J Obstet Gynecol* 1989; **161**(4): 996-1001.
3. Sato M, Noguchi J, Mashima M, Tanaka H, Hata T. 3D power Doppler ultrasound assessment of placental perfusion during uterine contraction in labor. *Placenta* 2016; **45**: 32-6.
4. Kiserud T, Acharya G. The fetal circulation. *Prenat Diagn* 2004; **24**(13): 1049-59.
5. Low JA. Intrapartum fetal asphyxia: definition, diagnosis, and classification. *Am J Obstet Gynecol* 1997; **176**(5): 957-9.
6. Ayres-de-Campos D, Arulkumaran S. FIGO consensus guidelines on intrapartum fetal monitoring: Physiology of fetal oxygenation and the main goals of intrapartum fetal monitoring. *Int J Gynaecol Obstet* 2015; **131**(1): 5-8.
7. Goldaber KG, Gilstrap LC, 3rd, Leveno KJ, Dax JS, McIntire DD. Pathologic fetal acidemia. *Obstet Gynecol* 1991; **78**(6): 1103-7.
8. Yeh P, Emary K, Impey L. The relationship between umbilical cord arterial pH and serious adverse neonatal outcome: analysis of 51,519 consecutive validated samples. *BJOG : an international journal of obstetrics and gynaecology* 2012; **119**(7): 824-31.
9. MacLennan A. A template for defining a causal relation between acute intrapartum events and cerebral palsy: international consensus statement. *Bmj* 1999; **319**(7216): 1054-9.
10. Sehdev HM, Stamilio DM, Macones GA, Graham E, Morgan MA. Predictive factors for neonatal morbidity in neonates with an umbilical arterial cord pH less than 7.00. *Am J Obstet Gynecol* 1997; **177**(5): 1030-4.
11. Wayenberg JL. Threshold of metabolic acidosis associated with neonatal encephalopathy in the term newborn. *J Matern Fetal Neonatal Med* 2005; **18**(6): 381-5.
12. Low JA, Lindsay BG, Derrick EJ. Threshold of metabolic acidosis associated with newborn complications. *Am J Obstet Gynecol* 1997; **177**(6): 1391-4.
13. Belfort MA, Saade GR, Thom E, et al. A Randomized Trial of Intrapartum Fetal ECG ST-Segment Analysis. *N Engl J Med* 2015; **373**(7): 632-41.
14. Brocklehurst P, Johns N, Johnston T, et al. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet (London, England)* 2017; **389**(10080): 1719-29.
15. Jonsson M, Soderling J, Ladfors L, et al. Implementation of a revised classification for intrapartum fetal heart rate monitoring and association to birth outcome: A national cohort study. *Acta obstetricia et gynecologica Scandinavica* 2022; **101**(2): 183-92.
16. Morales P, Bustamante D, Espina-Marchant P, et al. Pathophysiology of perinatal asphyxia: can we predict and improve individual outcomes? *Epma j* 2011; **2**(2): 211-30.

17. Rei M, Ayres-de-Campos D, Bernardes J. Neurological damage arising from intrapartum hypoxia/acidosis. *Best Pract Res Clin Obstet Gynaecol* 2016; **30**: 79-86.

18. Annual Report of the Swedish Pregnancy Register. 2021. https://www.medscinet.com/GR/uploads/hemsida/dokumentarkiv/Graviditetsregistrets%20%C3%85rsrapport%202021_1.0.pdf

19. ACOG Practice Bulletin No. 204: Fetal Growth Restriction. *Obstet Gynecol* 2019; **133**(2): e97-e109.

20. McIntire DD, Bloom SL, Casey BM, Leveno KJ. Birth weight in relation to morbidity and mortality among newborn infants. *N Engl J Med* 1999; **340**(16): 1234-8.

21. Battaglia FC, Lubchenco LO. A practical classification of newborn infants by weight and gestational age. *J Pediatr* 1967; **71**(2): 159-63.

22. Lees CC, Romero R, Stampalija T, et al. Clinical Opinion: The diagnosis and management of suspected fetal growth restriction: an evidence-based approach. *Am J Obstet Gynecol* 2022; **226**(3): 366-78.

23. McCowan LM, Figueras F, Anderson NH. Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy. *Am J Obstet Gynecol* 2018; **218**(2s): S855-s68.

24. Sharp A, Duong C, Agarwal U, Alfirevic Z. Screening and management of the small for gestational age fetus in the UK: A survey of practice. *European journal of obstetrics, gynecology, and reproductive biology* 2018; **231**: 220-4.

25. Vasak B, Koenen SV, Koster MP, et al. Human fetal growth is constrained below optimal for perinatal survival. *Ultrasound Obstet Gynecol* 2015; **45**(2): 162-7.

26. Gordijn SJ, Beune IM, Thilaganathan B, et al. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol* 2016; **48**(3): 333-9.

27. Kayem G, Grangé G, Bréart G, Goffinet F. Comparison of fundal height measurement and sonographically measured fetal abdominal circumference in the prediction of high and low birth weight at term. *Ultrasound Obstet Gynecol* 2009; **34**(5): 566-71.

28. Hon EH, Hess OW. Instrumentation of fetal electrocardiography. *Science* 1957; **125**(3247): 553-4.

29. SFOG. CTG och fosterövervakning Sweden. 2017. http://ctgutbildning.se.

30. Ayres-de-Campos D, Spong CY, Chandraharan E, Panel FIFMEC. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *Int J Gynaecol Obstet* 2015; **131**(1): 13-24.

31. Ghi T, Di Pasquo E, Dall'Asta A, et al. Intrapartum fetal heart rate between 150 and 160 bpm at or after 40 weeks and labor outcome. *Acta obstetricia et gynecologica Scandinavica* 2021; **100**(3): 548-54.

32. Serra V, Bellver J, Moulden M, Redman CW. Computerized analysis of normal fetal heart rate pattern throughout gestation. *Ultrasound Obstet Gynecol* 2009; **34**(1): 74-9.

33. Holzmann M, Wretler S, Cnattingius S, Nordstrom L. Cardiotocography patterns and risk of intrapartum fetal acidemia. *J Perinat Med* 2015; **43**(4): 473-9.

34. Cahill AG, Roehl KA, Odibo AO, Macones GA. Association and prediction of neonatal acidemia. *Am J Obstet Gynecol* 2012; **207**(3): 206 e1-8.

35. Triebwasser JE, Colvin R, Macones GA, Cahill AG. Nonreassuring Fetal Status in the Second Stage of Labor: Fetal Monitoring Features and Association with Neonatal Outcomes. *American journal of perinatology* 2016; **33**(7): 665-70.

36. Skupski DW, Rosenberg CR, Eglinton GS. Intrapartum fetal stimulation tests: a meta-analysis. *Obstet Gynecol* 2002; **99**(1): 129-34.

37. Holzmann M, Wretler S, Nordstrom L. Absence of accelerations during labor is of little value in interpreting fetal heart rate patterns. *Acta obstetricia et gynecologica Scandinavica* 2016; **95**(10): 1097-103.

38. Hon EH, Quilligan EJ. The classification of fetal heart rate. II. A revised working classification. *Conn Med* 1967; **31**(11): 779-84.

39. Lee ST, Hon EH. FETAL HEMODYNAMIC RESPONSE TO UMBILICAL CORD COMPRESSION. *Obstet Gynecol* 1963; **22**: 553-62.

40. Ball RH, Parer JT. The physiologic mechanisms of variable decelerations. *Am J Obstet Gynecol* 1992; **166**(6 Pt 1): 1683-8; discussion 8-9.

41. Lear CA, Galinsky R, Wassink G, et al. The myths and physiology surrounding intrapartum decelerations: the critical role of the peripheral chemoreflex. *J Physiol* 2016; **594**(17): 4711-25.

42. Lear CA, Westgate JA, Ugwumadu A, et al. Understanding Fetal Heart Rate Patterns That May Predict Antenatal and Intrapartum Neural Injury. *Semin Pediatr Neurol* 2018; **28**: 3-16.

43. Westgate JA, Bennet L, Gunn AJ. Fetal heart rate variability changes during brief repeated umbilical cord occlusion in near term fetal sheep. *British journal of obstetrics and gynaecology* 1999; **106**(7): 664-71.

44. Bennet L, Westgate JA, Liu YC, Wassink G, Gunn AJ. Fetal acidosis and hypotension during repeated umbilical cord occlusions are associated with enhanced chemoreflex responses in near-term fetal sheep. *J Appl Physiol (1985)* 2005; **99**(4): 1477-82.

45. Westgate JA, Bennet L, de Haan HH, Gunn AJ. Fetal heart rate overshoot during repeated umbilical cord occlusion in sheep. *Obstet Gynecol* 2001; **97**(3): 454-9.

46. Wassink G, Bennet L, Davidson JO, Westgate JA, Gunn AJ. Pre-existing hypoxia is associated with greater EEG suppression and early onset of evolving seizure activity during brief repeated asphyxia in near-term fetal sheep. *PLoS One* 2013; **8**(8): e73895.

47. Krebs HB, Petres RE, Dunn LJ. Intrapartum fetal heart rate monitoring. VIII. Atypical variable decelerations. *Am J Obstet Gynecol* 1983; **145**(3): 297-305.

48. National Institute for Health and Care Excellence: Guidelines. Intrapartum care for healthy women and babies. London: National Institute for Health and Care Excellence (NICE) Copyright © NICE 2020.; 2017.

49. Macones GA, Hankins GD, Spong CY, Hauth J, Moore T. The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. *Obstet Gynecol* 2008; **112**(3): 661-6.

50. Sholapurkar SL. Obstetrics at Decisive Crossroads Regarding Pattern-Recognition of Fetal Heart Rate Decelerations: Scientific Principles and Lessons From Memetics. *J Clin Med Res* 2018; **10**(4): 302-8.

51. Hamilton E, Warrick P, O'Keeffe D. Variable decelerations: do size and shape matter? *J Matern Fetal Neonatal Med* 2012; **25**(6): 648-53.

52. Marti Gamboa S, Pascual Mancho J, Lapresta Moros M, Rodriguez Lazaro L, de Bonrostro Torralba C, Castan Mateo S. Atypical decelerations: do they matter? *J Matern Fetal Neonatal Med* 2016: 1-5.

53. Shelley T, Tipton RH. Dip area. A quantitative measure of fetal heart rate patterns. *J Obstet Gynaecol Br Commonw* 1971; **78**(8): 694-701.

54. Martí Gamboa S, Lapresta Moros M, Pascual Mancho J, Lapresta Moros C, Castán Mateo S. Deceleration area and fetal acidemia. *J Matern Fetal Neonatal Med* 2017; **30**(21): 2578-84.

55. Giannubilo SR, Buscicchio G, Gentilucci L, Palla GP, Tranquilli AL. Deceleration area of fetal heart rate trace and fetal acidemia at delivery: a case-control study. *J Matern Fetal Neonatal Med* 2007; **20**(2): 141-4.

56. Strachan BK, Sahota DS, van Wijngaarden WJ, James DK, Chang AM. Computerised analysis of the fetal heart rate and relation to acidaemia at delivery. *BJOG : an international journal of obstetrics and gynaecology* 2001; **108**(8): 848-52.

57. Furukawa A, Neilson D, Hamilton E. Cumulative deceleration area: a simplified predictor of metabolic acidemia. *J Matern Fetal Neonatal Med* 2021; **34**(19): 3104-11.

58. Georgieva A, Redman CWG, Papageorghiou AT. Computerized data-driven interpretation of the intrapartum cardiotocogram: a cohort study. *Acta obstetricia et gynecologica Scandinavica* 2017; **96**(7): 883-91.

59. Georgieva A, Lear CA, Westgate JA, et al. Deceleration area and capacity during labour-like umbilical cord occlusions identify evolving hypotension: a controlled study in fetal sheep. *BJOG : an international journal of obstetrics and gynaecology* 2021; **128**(9): 1433-42.

60. Rivolta MW, Stampalija T, Casati D, et al. Acceleration and deceleration capacity of fetal heart rate in an in-vivo sheep model. *PLoS One* 2014; **9**(8): e104193.

61. Vintzileos AM, Smulian JC. Decelerations, tachycardia, and decreased variability: have we overlooked the significance of longitudinal fetal heart rate changes for detecting intrapartum fetal hypoxia? *Am J Obstet Gynecol* 2016; **215**(3): 261-4.

62. Chandraharan E. Intrapartum care: An urgent need to question historical practices and 'non-evidence'-based, illogical foetal monitoring guidelines to avoid patient harm. *Journal of Patient Safety and Risk Management* 2019; **24**(5): 210-7.

63. Hon EH. An Atlas of Fetal Heart Rate Patterns: Harty Press; 1968.

64. Alfirevic Z, Devane D, Gyte GM, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *The Cochrane database of systematic reviews* 2017; **2**: CD006066.

65. Public Health Agency of Sweden. Spädbarnsdödlighet [Internet]. Stockholm: Folkhälsomyndigheten; 2022. https://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/tolkad-rapportering/folkhalsans-utveckling/resultat/halsa/spadbarnsdodlighet/

66. Vintzileos AM, Antsaklis A, Varvarigos I, Papas C, Sofatzis I, Montgomery JT. A randomized trial of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation. *Obstet Gynecol* 1993; **81**(6): 899-907.

67. MacDonald D, Grant A, Sheridan-Pereira M, Boylan P, Chalmers I. The Dublin randomized controlled trial of intrapartum fetal heart rate monitoring. *Am J Obstet Gynecol* 1985; **152**(5): 524-39.

68. Palmer L, Blair E, Petterson B, Burton P. Antenatal antecedents of moderate and severe cerebral palsy. *Paediatr Perinat Epidemiol* 1995; **9**(2): 171-84.

69. Phelan JP, Ahn MO. Perinatal observations in forty-eight neurologically impaired term infants. *Am J Obstet Gynecol* 1994; **171**(2): 424-31.

70. Rhose S, Heinis AM, Vandenbussche F, van Drongelen J, van Dillen J. Inter- and intra-observer agreement of non-reassuring cardiotocography analysis and subsequent clinical management. *Acta obstetricia et gynecologica Scandinavica* 2014; **93**(6): 596-602.

71. Palomaki O, Luukkaala T, Luoto R, Tuimala R. Intrapartum cardiotocography -- the dilemma of interpretational variation. *J Perinat Med* 2006; **34**(4): 298-302.

72. 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**(6): 623 e1-5.

73. Blackwell SC, Grobman WA, Antoniewicz L, Hutchinson M, Gyamfi Bannerman C. Interobserver and intraobserver reliability of the NICHD 3-Tier Fetal Heart Rate Interpretation System. *Am J Obstet Gynecol* 2011; **205**(4): 378 e1-5.

74. Epstein AJ, Twogood S, Lee RH, Opper N, Beavis A, Miller DA. Interobserver reliability of fetal heart rate pattern interpretation using NICHD definitions. *American journal of perinatology* 2013; **30**(6): 463-8.

75. Nelson KB, Dambrosia JM, Ting TY, Grether JK. Uncertain value of electronic fetal monitoring in predicting cerebral palsy. *N Engl J Med* 1996; **334**(10): 613-8.

76. Jackson M, Holmgren CM, Esplin MS, Henry E, Varner MW. Frequency of fetal heart rate categories and short-term neonatal outcome. *Obstet Gynecol* 2011; **118**(4): 803-8.

77. Olofsson P, Ekengard F, Herbst A. Time to reconsider: Have the 2015 FIGO and 2017 Swedish intrapartum cardiotocogram classifications led us from Charybdis to Scylla? *Acta obstetricia et gynecologica Scandinavica* 2021; **100**(9): 1549-56.

78. Ingemarsson I, Arulkumaran S, Ingemarsson E, Tambyraja RL, Ratnam SS. Admission test: a screening test for fetal distress in labor. *Obstet Gynecol* 1986; **68**(6): 800-6.

79. Liston R, Sawchuck D, Young D. No. 197b-Fetal Health Surveillance: Intrapartum Consensus Guideline. *J Obstet Gynaecol Can* 2018; **40**(4): e298-e322.

80. Yli B KJ, Eikeland T, Henriksen T, Hjelle S, Blix E, et al. Fosterovervåking under fødsel, avnavling og syre-baseprøver fra navlesnor (Fetal surveillance during labour, cord clamping and cord blood gases). 2014.

81. WHO Guidelines Approved by the Guidelines Review Committee. WHO recommendations: Intrapartum care for a positive childbirth experience. Geneva: World Health Organization Copyright © World Health Organization 2018.; 2018.

82. Low JA, Pickersgill H, Killen H, Derrick EJ. The prediction and prevention of intrapartum fetal asphyxia in term pregnancies. *Am J Obstet Gynecol* 2001; **184**(4): 724-30.

83. SBU. Vetenskaplig evidens för användning av intagningstest med kardiotokografi (CTG) vid lågriskförlossningar. 2021. https://www.sbu.se/sv/publikationer/regional-

hta/vetenskaplig-evidens-for--anvandning-av-intagningstest--med-kardiotokografi-ctg-vid--lagriskforlossningar/

84. Impey L, Reynolds M, MacQuillan K, Gates S, Murphy J, Sheil O. Admission cardiotocography: a randomised controlled trial. *Lancet (London, England)* 2003; **361**(9356): 465-70.

85. Devane D, Lalor JG, Daly S, McGuire W, Cuthbert A, Smith V. Cardiotocography versus intermittent auscultation of fetal heart on admission to labour ward for assessment of fetal wellbeing. *The Cochrane database of systematic reviews* 2017; **1**: CD005122.

86. Smith V, Begley C, Newell J, et al. Admission cardiotocography versus intermittent auscultation of the fetal heart in low-risk pregnancy during evaluation for possible labour admission - a multicentre randomised trial: the ADCAR trial. *BJOG : an international journal of obstetrics and gynaecology* 2019; **126**(1): 114-21.

87. Epplin KA, Tuuli MG, Odibo AO, Roehl KA, Macones GA, Cahill AG. Effect of Growth Restriction on Fetal Heart Rate Patterns in the Second Stage of Labor. *American journal of perinatology* 2015; **32**(9): 873-8.

88. Dawes GS, Moulden M, Redman CW. System 8000: computerized antenatal FHR analysis. *J Perinat Med* 1991; **19**(1-2): 47-51.

89. Dawes GS, Moulden M, Redman CW. Short-term fetal heart rate variation, decelerations, and umbilical flow velocity waveforms before labor. *Obstet Gynecol* 1992; **80**(4): 673-8.

90. Serra V, Moulden M, Bellver J, Redman CW. The value of the short-term fetal heart rate variation for timing the delivery of growth-retarded fetuses. *BJOG : an international journal of obstetrics and gynaecology* 2008; **115**(9): 1101-7.

91. Nunes I, Ayres-de-Campos D, Ugwumadu A, et al. Central Fetal Monitoring With and Without Computer Analysis A Randomized Controlled Trial. *Obstetrics and Gynecology* 2017; **129**(1): 83-90.

92. Ignatov PN, Lutomski JE. Quantitative cardiotocography to improve fetal assessment during labor: a preliminary randomized controlled trial. *European journal of obstetrics, gynecology, and reproductive biology* 2016; **205**: 91-7.

93. Campanile M, D'Alessandro P, Della Corte L, et al. Intrapartum cardiotocography with and without computer analysis: a systematic review and meta-analysis of randomized controlled trials. *J Matern Fetal Neonatal Med* 2020; **33**(13): 2284-90.

94. Iraji MS. Prediction of fetal state from the cardiotocogram recordings using neural network models. *Artif Intell Med* 2019; **96**: 33-44.

95. Comert Z, Sengur A, Budak U, Kocamaz AF. Prediction of intrapartum fetal hypoxia considering feature selection algorithms and machine learning models. *Health Inf Sci Syst* 2019; **7**(1): 17.

96. Saling E. Amnioscopy and foetal blood sampling: observations on foetal acidosis. *Arch Dis Child* 1966; **41**(219): 472-6.

97. Nordstrom L, Ingemarsson I, Kublickas M, Persson B, Shimojo N, Westgren M. Scalp blood lactate: a new test strip method for monitoring fetal wellbeing in labour. *British journal of obstetrics and gynaecology* 1995; **102**(11): 894-9.

98. East CE, Leader LR, Sheehan P, Henshall NE, Colditz PB, Lau R. Intrapartum fetal scalp lactate sampling for fetal assessment in the presence of a non-reassuring fetal heart rate trace. *The Cochrane database of systematic reviews* 2015; **5**: CD006174.

99. Wiberg-Itzel E, Lipponer C, Norman M, et al. Determination of pH or lactate in fetal scalp blood in management of intrapartum fetal distress: randomised controlled multicentre trial. *BMJ* 2008; **336**(7656): 1284-7.

100. Westgren M, Kruger K, Ek S, et al. Lactate compared with pH analysis at fetal scalp blood sampling: a prospective randomised study. *British journal of obstetrics and gynaecology* 1998; **105**(1): 29-33.

101. Nordstrom L, Ingemarsson I, Westgren M. Fetal monitoring with lactate. *Baillieres Clin Obstet Gynaecol* 1996; **10**(2): 225-42.

102. Kruger K, Hallberg B, Blennow M, Kublickas M, Westgren M. Predictive value of fetal scalp blood lactate concentration and pH as markers of neurologic disability. *Am J Obstet Gynecol* 1999; **181**(5 Pt 1): 1072-8.

103. Birgisdottir BT, Holzmann M, Varli IH, Graner S, Saltvedt S, Nordström L. Reference values for Lactate Pro 2™ in fetal blood sampling during labor: a cross-sectional study. *J Perinat Med* 2017; **45**(3): 321-5.

104. Jorgensen JS, Weber T. Fetal scalp blood sampling in labor--a review. *Acta obstetricia et gynecologica Scandinavica* 2014; **93**(6): 548-55.

105. East CE, Leader LR, Sheehan P, Henshall NE, Colditz PB, Lau R. Intrapartum fetal scalp lactate sampling for fetal assessment in the presence of a non-reassuring fetal heart rate trace. *The Cochrane database of systematic reviews* 2015; (5): CD006174.

106. East CE, Davey MA, Kamlin COF, et al. The addition of fetal scalp blood lactate measurement as an adjunct to cardiotocography to reduce caesarean sections during labour: The Flamingo randomised controlled trial. *Aust N Z J Obstet Gynaecol* 2021; **61**(5): 684-92.

107. East CE, Begg L, Colditz PB, Lau R. Fetal pulse oximetry for fetal assessment in labour. *The Cochrane database of systematic reviews* 2014; (10): CD004075.

108. Rosen KG. Fetal electrocardiogram waveform analysis in labour. *Curr Opin Obstet Gynecol* 2005; **17**(2): 147-50.

109. Saccone G, Schuit E, Amer-Wahlin I, Xodo S, Berghella V. Electrocardiogram ST Analysis During Labor: A Systematic Review and Meta-analysis of Randomized Controlled Trials. *Obstet Gynecol* 2016; **127**(1): 127-35.

110. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; **85**(3): 257-68.

111. Lu K, Holzmann M, Abtahi F, Lindecrantz K, Lindqvist PG, Nordstrom L. Fetal heart rate short term variation during labor in relation to scalp blood lactate concentration. *Acta obstetricia et gynecologica Scandinavica* 2018; **97**(10): 1274-80.

112. Westerhuis ME, van Horen E, Kwee A, van der Tweel I, Visser GH, Moons KG. Inter- and intra-observer agreement of intrapartum ST analysis of the fetal electrocardiogram in women monitored by STAN. *BJOG : an international journal of obstetrics and gynaecology* 2009; **116**(4): 545-51.

113. Blix E. The admission CTG: is there any evidence for still using the test? *Acta obstetricia et gynecologica Scandinavica* 2013; **92**(6): 613-9.

114. 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**(7300): 1457-60; discussion 60-2.

115. Blix E, Sviggum O, Koss KS, Oian P. Inter-observer variation in assessment of 845 labour admission tests: comparison between midwives and obstetricians in the clinical setting and two experts. *BJOG : an international journal of obstetrics and gynaecology* 2003; **110**(1): 1-5.

116. 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**(1): 33-7.

117. Ayres-de-Campos D, Bernardes J. Early, variable and late decelerations: can a consensus be reached in their identification? *Int J Gynaecol Obstet* 1999; **65**(3): 305-6.

118. Nunes I, Ayres-de-Campos D. Computer analysis of foetal monitoring signals. *Best Pract Res Clin Obstet Gynaecol* 2016; **30**: 68-78.

119. Cahill AG, Tuuli MG, Stout MJ, Lopez JD, Macones GA. A prospective cohort study of fetal heart rate monitoring: deceleration area is predictive of fetal acidemia. *Am J Obstet Gynecol* 2018; **218**(5): 523 e1- e12.

120. Graham EM, Adami RR, McKenney SL, Jennings JM, Burd I, Witter FR. Diagnostic accuracy of fetal heart rate monitoring in the identification of neonatal encephalopathy. *Obstet Gynecol* 2014; **124**(3): 507-13.

121. Westgate JA, Wibbens B, Bennet L, Wassink G, Parer JT, Gunn AJ. The intrapartum deceleration in center stage: a physiologic approach to the interpretation of fetal heart rate changes in labor. *Am J Obstet Gynecol* 2007; **197**(3): 236 e1-11.

122. Garite TJ, Dildy GA, McNamara H, et al. A multicenter controlled trial of fetal pulse oximetry in the intrapartum management of nonreassuring fetal heart rate patterns. *Am J Obstet Gynecol* 2000; **183**(5): 1049-58.

123. Kubli FW, Hon EH, Khazin AF, Takemura H. Observations on heart rate and pH in the human fetus during labor. *Am J Obstet Gynecol* 1969; **104**(8): 1190-206.

124. Liston WA. Admission cardiotocography. *Lancet (London, England)* 2003; **361**(9370): 1741; author reply -2.

125. Chalmers I. Cardiotocography v Doppler auscultation. All unbiased comparative studies should be published. *Bmj* 2002; **324**(7335): 483-5.

126. Grant S. Cardiotocography v Doppler auscultation. Conclusions do not recognise difference between statistical and clinical significance. *Bmj* 2002; **324**(7335): 483.

127. RANZCOG. Intrapartum Fetal Surveillance. 2019. https://ranzcog.edu.au/wp-content/uploads/2022/05/Intrapartum-Fetal-Surveillance.pdf.

128. Rosset IK, Lindahl K, Blix E, Kaasen A. Recommendations for intrapartum fetal monitoring are not followed in low-risk women: A study from two Norwegian birth units. *Sexual & reproductive healthcare : official journal of the Swedish Association of Midwives* 2020; **26**: 100552.

129. LÖF 2015. https://lof.se/filer/Fosterovervakning-vid-aktiv-forlossning.pdf.

130. Lindqvist PG, Molin J. Does antenatal identification of small-for-gestational age fetuses significantly improve their outcome? *Ultrasound Obstet Gynecol* 2005; **25**(3): 258-64.

131. Chauhan SP, Rice MM, Grobman WA, et al. Neonatal Morbidity of Small- and Large-for-Gestational-Age Neonates Born at Term in Uncomplicated Pregnancies. *Obstet Gynecol* 2017; **130**(3): 511-9.

132. Akhavan S, Lak P, Rahimi-Sharbaf F, Mohammadi SR, Shirazi M. Admission Test and Pregnancy Outcome. *Iran J Med Sci* 2017; **42**(4): 362-8.

133. Lovers AAK, Ugwumadu A, Georgieva A. Cardiotocography and Clinical Risk Factors in Early Term Labor: A Retrospective Cohort Study Using Computerized Analysis With Oxford System. *Front Pediatr* 2022; **10**: 784439.

134. Sholapurkar SL. Intermittent Auscultation in Labor: Could It Be Missing Many Pathological (Late) Fetal Heart Rate Decelerations? Analytical Review and Rationale for Improvement Supported by Clinical Cases. *J Clin Med Res* 2015; **7**(12): 919-25.

135. Sholapurkar SL. Intermittent auscultation (surveillance) of fetal heart rate in labor: a progressive evidence-backed approach with aim to improve methodology, reliability and safety. *J Matern Fetal Neonatal Med* 2022; **35**(15): 2942-8.

136. Costa MA, Ayres-de-Campos D, Machado AP, Santos CC, Bernardes J. Comparison of a computer system evaluation of intrapartum cardiotocographic events and a consensus of clinicians. *J Perinat Med* 2010; **38**(2): 191-5.

137. Parer JT, Hamilton EF. Comparison of 5 experts and computer analysis in rule-based fetal heart rate interpretation. *Am J Obstet Gynecol* 2010; **203**(5): 451 e1-7.

138. Arias F. Intrauterine resuscitation with terbutaline: a method for the management of acute intrapartum fetal distress. *Am J Obstet Gynecol* 1978; **131**(1): 39-43.

139. Tejani NA, Verma UL, Chatterjee S, Mittelmann S. Terbutaline in the management of acute intrapartum fetal acidosis. *J Reprod Med* 1983; **28**(12): 857-61.