Fetal electrocardiography ST-segment analysis for intrapartum monitoring: a critical appraisal of conflicting evidence and a way forward

Isis Amer-Wåhlin, MD, PhD; Austin Ugwumadu, MD, PhD; Branka M. Yli, MD, PhD; Anneke Kwee, MD, PhD; Susanna Timonen, MD, PhD; Vanessa Cole, RM, MSc; Diogo Ayres-de-Campos, MD, PhD; Georges-Emmanuel Roth, RM, MSc; Christiane Schwarz, RM, PhD; Luca A. Ramenghi, MD, PhD; Tullia Todros, MD, PhD; Virginie Ehlinger, MS; Christophe Vayssiere, MD, PhD; On behalf of the Study Group of Intrapartum Fetal Monitoring (European Association of Perinatal Medicine)

BACKGROUND: In the past century, some areas of obstetric including intrapartum care have been slow to benefit from the dramatic advances in technology and medical care. Although fetal heart rate monitoring (cardiotocography) became available a half century ago, its interpretation often differs between institutions and countries, its diagnostic accuracy needs improvement, and a technology to help reduce the unnecessary obstetric interventions that have accompanied the cardiotocography is urgently needed.

STUDY DESIGN: During the second half of the 20th century, key findings in animal experiments captured the close relationship between myocardial glycogenolysis, myocardial workload, and ST changes, thus demonstrating that ST waveform analysis of the fetal electrocardiogram can provide information on oxygenation of the fetal myocardium and establishing the physiological basis for the use of electrocardiogram in intrapartum fetal surveillance.

RESULTS: Six randomized controlled trials, 10 meta-analyses, and more than 20 observational studies have evaluated the technology developed based on this principle. Nonetheless, despite this intensive assessment, differences in study protocols, inclusion criteria, enrollment rates, clinical guidelines, use of fetal blood sampling, and definitions of key outcome parameters, as well as inconsistencies in randomized controlled trial data handling and statistical methodology, have made this voluminous evidence difficult to interpret. Enormous resources spent on randomized controlled trials have failed to guarantee the generalizability of their results to other settings or their ability to reflect everyday clinical practice.

CONCLUSION: The latest meta-analysis used revised data from primary randomized controlled trials and data from the largest randomized controlled trials from the United States to demonstrate a significant reduction of metabolic acidosis rates by 36% (odds ratio, 0.64; 95% confidence interval, 0.46–0.88) and operative vaginal delivery rates by 8% (relative risk, 0.92; 95% confidence interval, 0.86–0.99), compared with cardiotocography alone.

Key words: acidemia, cardiotocography, cesarean delivery, fetal blood sampling, fetal death, fetal distress, fetal electrocardiogram, fetal heart rate, intrapartum fetal monitoring, meta-analysis, observational studies, perinatal asphyxia, randomized controlled trials, real-world evidence, ST segment analysis

From the Medical Management Centre, Department of Learning, Informatics, Management and Ethics, Department of Women and Child Health, Karolinska Institute, Stockholm, Sweden (Dr Amer-Wåhlin); Department of Obstetrics & Gynaecology, St George's, University of London, London, UK (Dr Ugwumadu); Department of Obstetrics, Rikshospitalet Oslo University Hospital, Oslo (Dr Yli); Department of Obstetrics and Gynecology, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands (Dr Kwee); Department of Obstetrics and Gynecology, Turku University Hospital, Turku, Finland (Dr Timonen); Department of Obstetrics and Gynaecology, Kingston Hospital, Galsworthy Road, Kingston-upon-Thames, UK (Dr Cole); Department of Obstetrics and Gynecology, and Reproductive Medicine, Santa Maria Hospital, University of Lisbon School of Medicine, Lisbon, Portugal (Dr Ayres-de-Campos); Department of Obstetrics and Gynecology, CHRU of Strasbourg, Louis Pasteur University, Strasbourg, France (Dr Roth); Lübeck University, Midwifery Research and Education Unit, Department of Obstetrics, Gynaecology & Reproductive Medicine, Lübeck Germany (Dr Schwarz); Neonatal Intensive Care Unit, IRCCS Istituto Giannina Gaslini, Genova, Italy (Dr Ramenghi); Department of Obstetrics and Gynecology, Sant'Anna Hospital, University of Turin, Turin, Italy (Dr Todros); UMR 1027 INSERM, Team SPHERE, Toulouse III University, Toulouse, France (Drs Ehlinger and Vayssière); Department of Obstetrics and Gynaecology, Paule de Viguier Hospital, CHU Toulouse, Toulouse, France (Dr Vayssière).

Received Oct. 26, 2018; revised March 31, 2019; accepted April 1, 2019.

The authors report no conflict of interest.

Corresponding author: Isis Amer-Wåhlin, MD, PhD. isis.amer-wahlin@ki.se

0002-9378/$36.00 • © 2019 Published by Elsevier Inc. • https://doi.org/10.1016/j.ajog.2019.04.003

The purpose of fetal monitoring during labor is to predict significant intrapartum fetal hypoxia/acidosis and institute timely intervention to prevent injury without unnecessary operative delivery. Although the last 50 years have seen dramatic developments in technology and medical care, some areas of obstetrics including intrapartum care have been slow to benefit from these advances.

Since the work of Hon and Lee[1] and Caldeyro-Barcia[2] in the 1950s and the development of the first commercially available fetal heart rate monitors by Hammacher[3] and Hewlett Packard in 1968, progress in engineering and computer technology has led to improved signal quality and accuracy and thus made fetal monitoring standard of care. Yet the debate around the specificity and effectiveness of electronic fetal monitoring (currently considered a misnomer) for identifying fetal compromise by cardiotocography (CTG) and its impact on cesarean delivery rates and neonatal outcome remains intense.[4]

The increase in operative delivery associated with CTG has prompted questions around the definition of abnormal CTG patterns,[5] and some countries have issued recommendations to use only intermittent auscultation in low-risk labors.[6,7] CTG interpretation has an interobserver variability of 43% (283 of 660) between experts according to the International Federation of Gynecology and Obstetrics classification[8] and limited specificity for detecting fetal hypoxia/acidosis.

Interpretation guidelines have been developed, but these often differ between institutions and countries.[9,10] Few studies have examined the clinical impact of guidelines updated more recently,[11,12] both for CTG and new technologies added to CTG, such as fetal electrocardiogram (ECG) ST segment analysis (STAN; Neoventa Medical, Gothenburg, Sweden).[13]

While many centers have abandoned the use of an invasive fetal blood sampling (FBS) because of its poor predictive value and its rare but serious risks,[14] some northern European centers still use this procedure to assess pH or lactate to evaluate fetal oxygenation when nonreassuring CTG patterns occur, with the aim of minimizing unnecessary interventions.[14,15] However, few studies have validated the correlation between fetal blood sampling and perinatal hypoxia/acidosis,[16-18] and it remains uncertain whether it really decreases operative deliveries.[19] Rare complications with the technique have been reported, such as cerebrospinal fluid leakage, scalp abscesses, and fetal hemorrhage,[20,21] and many countries, including the United States, abandoned this technique some years ago.

Fetal scalp stimulation has also been used to assess fetal oxygenation,[22] but failure to respond to it with an acceleration does not identify a fetus with hypoxia/acidosis.[23] A meta-analysis concluded that FBS and scalp stimulation have similar predictive values for the prediction of acidemia at birth with negative likelihood ratios of, respectively, 0.12 (0.02−0.78) and 0.06 (0.01−0.31) and positive likelihood ratios of 8.54 (1.28−56.96) and 15.68 (3.22−76.24).[24] However, the data concerning these intermittent second line tests on hypoxia during labor are scarce and need to be confirmed.

For all these reasons, the utilization of the different types of intrapartum fetal monitoring with or without FBS or scalp stimulation remains variable between clinicians, hospitals, countries, and regions.

On the other hand, epidemiological data from the United States show a clear association between CTG monitoring and lower infant mortality.[25,26] The near elimination of intrapartum stillbirths in high-resource countries, attributed to the widespread use of intrapartum CTG,[27] suggests that intrapartum causes may be more common than previously suggested.[28] Several Scandinavian reports indicate that birth asphyxia is the likely cause of cerebral palsy (CP) in 28% of children with CP born at term.[29,30]

Despite the marked variations observed in CP rates in term newborns caused by perinatal factors, hypoxic-ischemic encephalopathy is the dominant clinical finding in most cases, 79% in these 2 Nordic registries. Moreover, Cowan et al[31] have reported in more than 90% of the 351 term infants with neonatal encephalopathy, seizures, or both, but without specific syndromes or major congenital defects, evidence of perinatal insults with brain magnetic resonance imaging or postmortem examination and a very low rate of brain injury acquired before birth. Even if the literature is limited, we may conclude that reducing the incidence of neonatal encephalopathy could potentially impact the long-term outcome in term newborns.

In this narrative review, we examine the physiological basis and development of ST-segment analysis of the fetal electrocardiogram, critically appraise the results of randomized clinical trials, systematic reviews, and meta-analyses of this technology, and propose an interpretation of the results obtained so far as well as a way forward for clinical implementation and further research.

The development of the fetal electrocardiogram

From the electrocardiogram to the fetal electrocardiogram

As decribed by Hurst,[32] toward the end of the 18th century, Waller and Einthoven identified the 5 deflections of Rosen and the electrocardiogram (PQRST) and described it with the nomenclature still used today.[32] For decades, assessment of ST interval deviations has been the cornerstone of diagnosis of ischemic heart disease.[33] Significant ST depression at peak exercise is a sign of myocardial ischemia, while in patients with chest pain, ST elevation is a specific sign of myocardial infarction.[34] Interestingly, T-wave amplitude also increases rapidly in healthy men subjected to hard anaerobic work and returns to normal after 8 minutes of rest.[35]

During the second half of the 20th century, it was postulated that ST waveform analysis of the fetal ECG could also provide information on oxygenation of the fetal myocardium. In 1957, Southern[36] reported the first comprehensive study of fetal ECG obtained with abdominal leads and related to fetal distress and arterial oxygen saturation at birth. Decreased umbilical cord blood

oxygen saturation at delivery was found to be associated with prolonged P-R, QRS, and ST intervals and with inversion of the T wave.[37,38] Poor signal-to-noise ratios, because of fetal ECG acquisition via abdominal electrodes and limited knowledge of the pathophysiology of fetal cardiac function during hypoxia/acidosis nonetheless delayed the clinical utility of these findings.

Signal quality improved with the introduction of fetal scalp electrodes and computer processing and led to the understanding of the relationship between ECG features and fetal acid-base status.[39,40] Prolongation of the QRS interval was found to be associated with baseline tachycardia, shortening of the P-R interval with CTG decelerations, and T-wave inversion with signs of placental dysfunction.[41] Pardi et al[42] showed changes in ECG configuration during periods of fetal hypoxia, with P-wave and P-Q interval modifications appearing consistently during late decelerations; ST depression/elevation/inversion or increased T-wave amplitude occurred in 48% of cases.

Fetal electrocardiogram and fetal oxygenation

By the mid-1970s, it was widely, although not universally,[38] believed that ECG changes were preceded by marked CTG changes.[43,44] The pioneering work of Rosen and Kjellmer[45] on the ECG analysis helped to elucidate the pathophysiological mechanisms involved in CTG changes and thereby enabled more precise evaluation of fetal oxygenation. It was not fully appreciated at the time that identification of ST waveform changes required faster signal processing capacities than CTG did to obtain a high-quality ECG or that the fetal scalp electrode needed ECG signal filtering features to remove the distortions of low-frequency ST changes. Furthermore, the optimal configuration of the fetal ECG lead was ultimately required to detect ST changes.[46]

Experimental studies

Animal experiments in rhesus monkeys demonstrated the possibility of evaluating fetal ST changes in intrapartum surveillance.[46] The severity of the ST-T changes in guinea pig fetuses correlated linearly and strongly with the depletion of myocardial glycogen and creatine-P; cellular adenosine triphosphate production remained unaffected until near-total glycogen depletion.[47]

At the same time, the myocardial glycogenolysis rate in lamb fetuses was linearly correlated with the increase in T-wave amplitude.[47] These key findings demonstrated the close relationship between myocardial glycogenolysis, the myocardial workload, and ST changes and thus established the physiological basis for the use of ECG in intrapartum fetal surveillance.

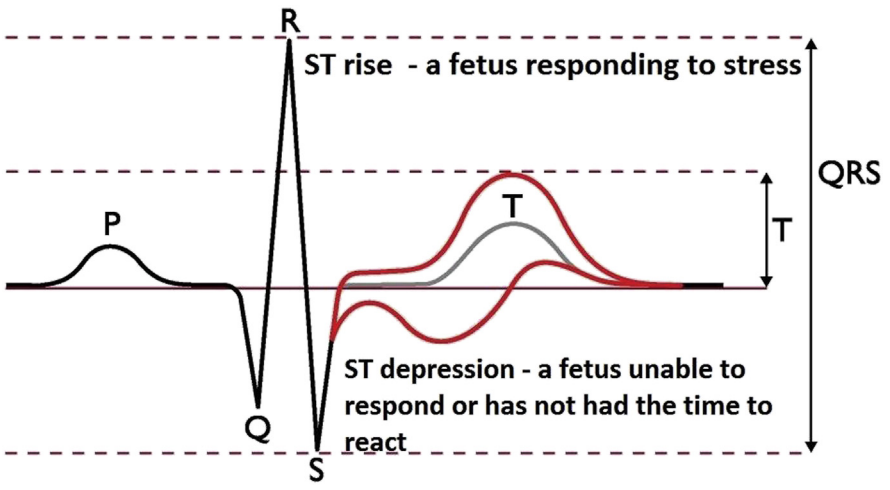
Pathophysiology underlying ST interval changes

The ST interval comprises the ST segment and the T wave, and both relate to the repolarization of myocardial cells in preparation for the next contraction, an energy-intensive process. An increase in T-wave height, quantified by the T/QRS ratio (Figure 1), occurs when cellular energy production within myocardial cells begins to decline,[45,48] that is, when the oxygen supply is inadequate to maintain metabolic activity so that cells are forced to generate energy by ß-adrenoceptor-mediated anaerobic breakdown of glycogen reserves.[49]

An important aspect of fetal response to hypoxia is a marked catecholamine surge. Figure 2 shows the relation between catecholamine concentrations and T/QRS ratios according to the severity of acidemia. Moderate hypoxemia increases T-wave amplitude, adrenaline surge, ß-adrenoceptor activation, and myocardial glycogenolysis. Potassium ions (K^+) released during myocardial glycogenolysis affect the myocardial cell membrane potential, causing a rise in the ST waveform and T-wave amplitude (Figure 1). This increase in the T-wave amplitude and the T/QRS ratio thus reflect the rate of myocardial glycogenolysis and the utilization of a vital compensatory defense mechanism to hypoxia.[50]

ST interval depression indicates an imbalance between the endocardium and epicardium because of the difference between the lower blood perfusion pressure of the endocardium and the higher mechanical strain, which delays myocardial repolarization. Unless the myocardium is generally activated (β-receptor activation and enhanced Frank-Starling relationship (the energy liberated with each contraction of the heart is a function of the length of the

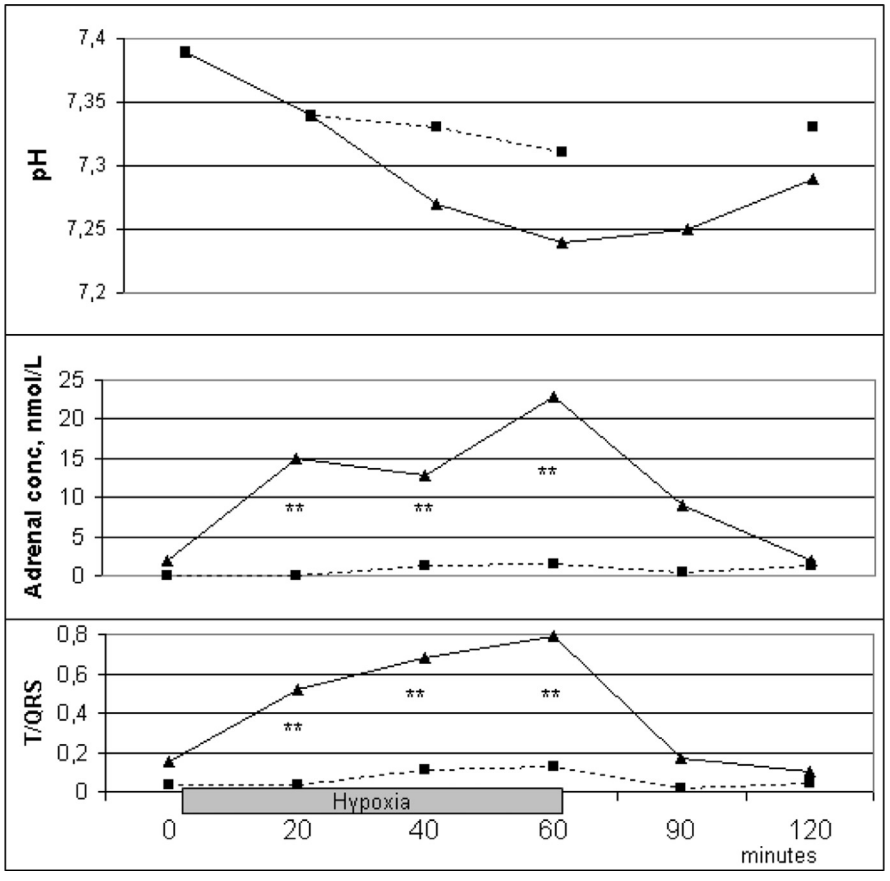
FIGURE 1
ECG waveform showing hypoxia-related ST segment changes



The T/QRS ratio measurement is also indicated (courtesy of Professor K. G. Rosen).

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

FIGURE 2
Changes in arterial pH, plasma adrenalin concentration, and T/QRS ratios



Changes in artery pH, plasma adrenalin concentration, and T/QRS of fetal lambs in connection with 60 minutes of hypoxia induced by letting the ewe breathe a 7% oxygen gas mixture. *Solid lines* indicate the group of fetuses (n = 10) that responded with an increase in T-wave amplitude, and *dashed lines* indicate group of fetuses with no ST change in response to hypoxia. The level of significance between the groups is indicated. *Double asterisk* indicates $P < .01$.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

myocardial cell fibers; preload, end-diastolic volume [or pressure]. Any decrease in performance may cause ST interval depression.

All the factors that modify the performance characteristics of the myocardial wall, including hypoxia, prematurity, infections, maternal fever, myocardial dystrophy, maternal diabetes,[51] and cardiac malformations,[52] may depress the ST interval (Figure 3). The ST patterns associated with cord occlusion and endotoxin administration observed in midgestation fetal lambs demonstrate the relationship between gestational age and response to hypoxia.[53]

During prolonged cord occlusion, blood pressure and ST waveform amplitude increase significantly, subsiding as hypoxia evolves.[53] Depressed ST intervals appear to coincide with cardiac dysfunction and hypotension. Responses to bacterial endotoxin administration develop over several hours and include hypotension, tachycardia, and ST depression.[54] Figure 4 illustrates a case of biphasic ST, which is the term used to describe depressed ST intervals.

Fetal electrocardiogram waveform analysis

Figure 1 shows the different parts of the ECG waveform. While the P and QRS waveforms are stable and easy to detect, the ST waveform changes are affected by signal noise and the mode of ECG recording. For example, in a fetal breech presentation, the P/QRS ratios will be inverted compared to a cephalic presentation.

The development of fetal ST analysis began with the intrapartum recording by a single-channel ECG recorder in June 1981.[55] A series of recordings using different microprocessor-based ST analyzers followed (Figure 5)[56] and made it evident that clinical use requires that it be combined with CTG analysis.[57] The availability of affordable processors in the 1990s enabled the development of signal processing and fetal ECG signal filters, which in turn allowed low-frequency noise reduction, consistent ST signal acquisition, and the introduction of automatic detection of relevant ST changes.[58]

Because averaging over 30 beats allowed noise reduction and detailed measurements of the ST waveform, it was considered acceptable. Signal processing algorithms were developed to detect changes in the ST segment slope (biphasic STs) and in the T/QRS amplitude. Biphasic STs were classified according to the relationship between the baseline and their slope into category 1 (slope above baseline), 2 (cross the baseline), or 3 (below baseline) (Figure 4). A T/QRS signal baseline was estimated within 4–5 minutes at system start-up (based on 20 accepted fetal ECGs) and used to detect subsequent T/QRS elevations, calculated by comparing each with the 10 most recent values.

Three types of relevant ST events were recognized and incorporated into signal processing algorithms: episodic T/QRS rise (T/QRS rise >0.10 in 2 consecutive T/QRSs); baseline T/QRS rise (T/QRS rise >0.05 in more than 10 minutes) (Figure 6); and ST interval depressions: categories 2 and 3 biphasic ST events. All these 3 types of ST events were associated with acidosis as was shown in experimental studies with appropriate cutoff for them to intervene in link with the International

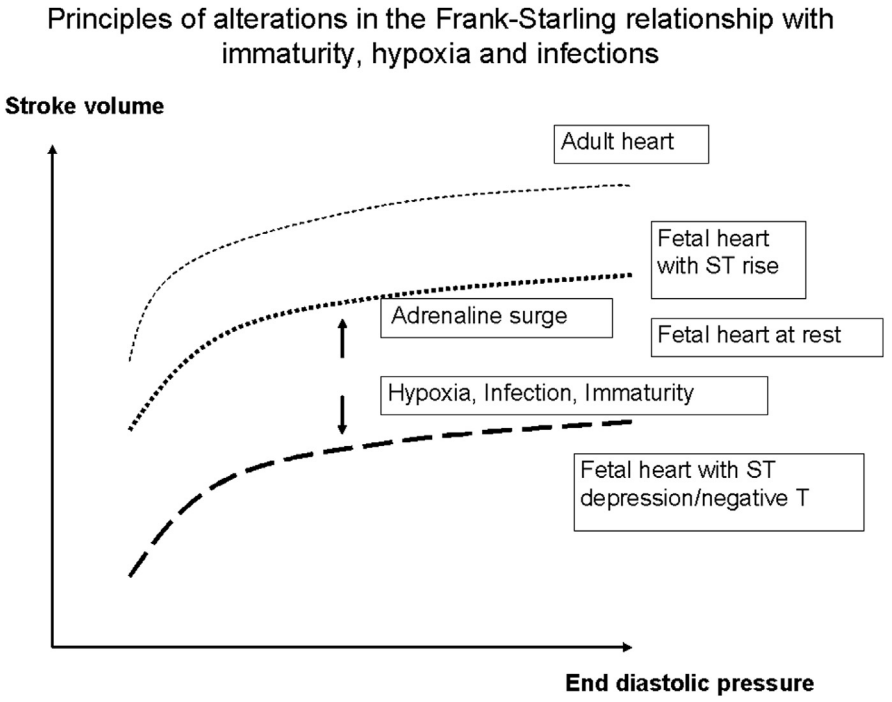
Federation of Gynecology and Obstetrics CTG classification (Figure 7).[52]

Randomized controlled trials and systematic reviews of ST analysis

After initial exploratory clinical studies and technological developments, clinical trials were conducted to determine whether the use of the ST waveform analysis in combination with CTG improved neonatal and maternal outcomes in high-risk pregnancies. Between 1993 and 2010, 5 randomized clinical trials were conducted in Europe.[59-63] The inconsistent results of these studies were evaluated using meta-analysis techniques in 6 publications between 2012 and 2014.[64-69] However, these meta-analyses and their results differed in many ways, causing a debate about the effectiveness of STAN.

In 2015, the results of a sixth randomized clinical trial performed in the United States were published.[70] Three meta-analyses integrating data from the 6 trials conducted since 1993 have since been published.[71-73] The choice of metabolic acidosis as outcome measure for the evaluation of the effectiveness of ST analysis is also a source of debate.

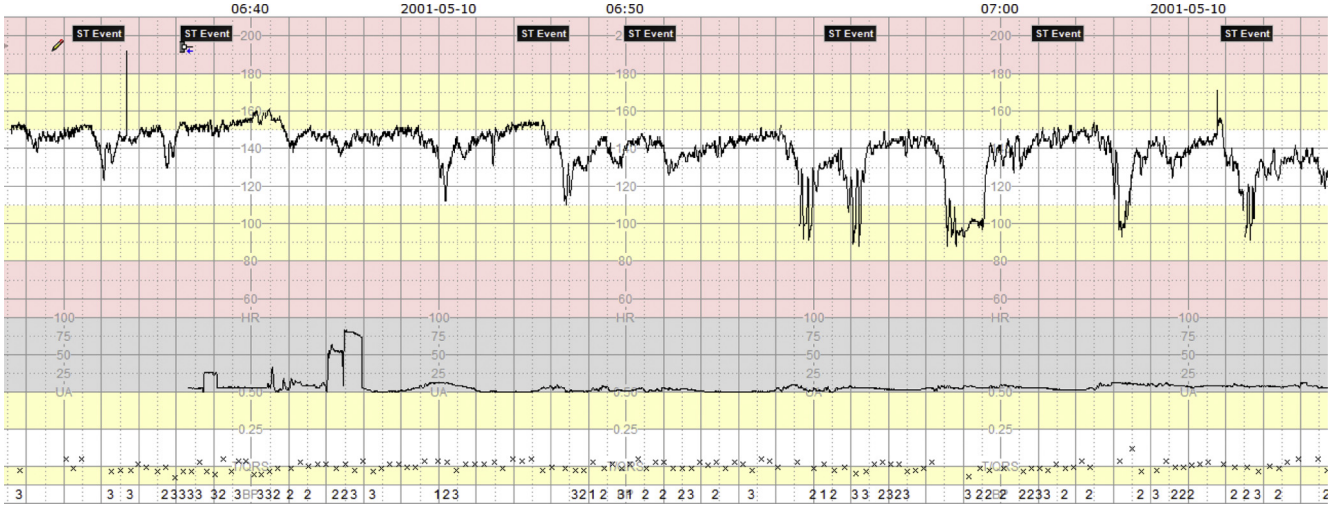
FIGURE 3
Changes in Frank-Starling relationship associated with different ST patterns



Illustrations of changes in the Frank-Starling relationship associated with different ST patterns (courtesy of Professor K. G. Rosén).

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

FIGURE 4
Example of STAN recording during active labor with biphasic ST events

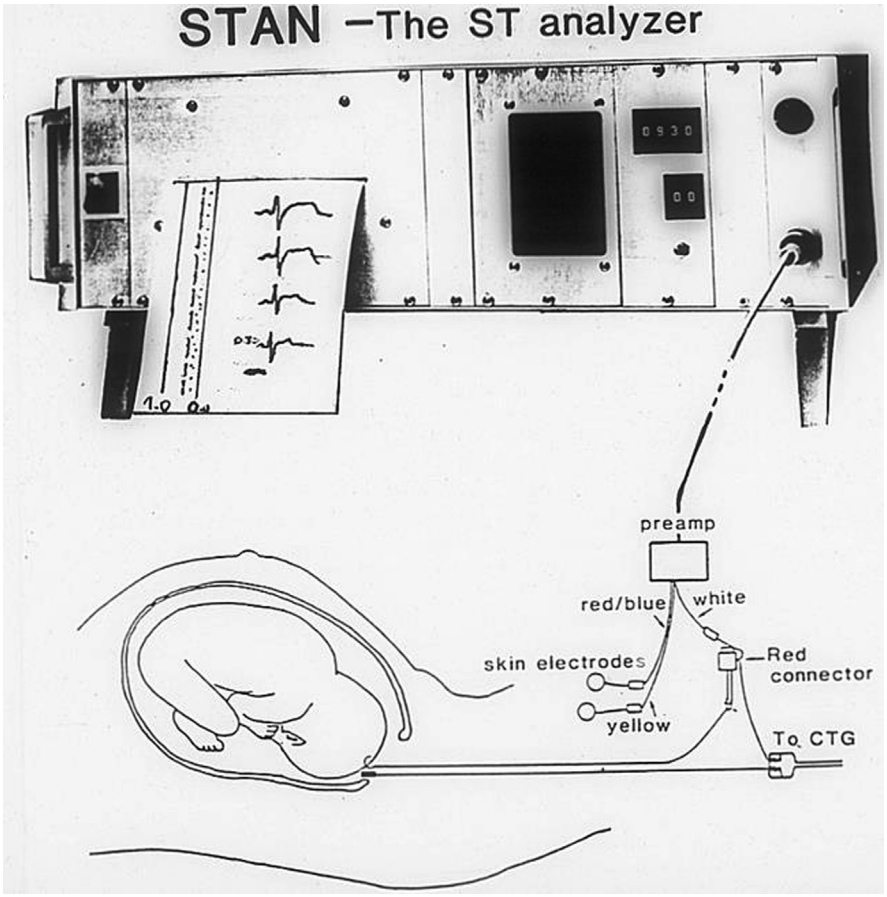


Note the occurrence of biphasic ST changes identified by 2 and 3. The corresponding fetal ECG complexes are also displayed. This case subsequently developed an ST rise and severe asphyxia.

ECG, electrocardiogram.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

FIGURE 5
One-channel ECG recorder



(Mk III dated 1981, courtesy of Professor K. G. Rosén).

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Selection of an appropriate outcome measure

Assessing the effectiveness of a clinical intervention requires the availability of a reasonably prevalent and specific marker of adverse outcome. Outcome measures such as perinatal mortality, CP, and neonatal encephalopathy are too rare in high-resource countries to be useful for this purpose.

Low umbilical artery pH is significantly associated with neonatal encephalopathy, intraventricular hemorrhage, periventricular leukomalacia, and CP.[74] Newborn umbilical artery pH <7.00 and its combination with base deficit (BD) >12 mmol/L (metabolic acidosis) have emerged as better predictors of severe neonatal morbidity and mortality.[74,75]

These criteria identify all neonates at risk of severe complications because of peripartum hypoxia, including many who recover quickly and have no long-term consequences.[74-76] Thus, the clinical trials that have been carried out have mainly chosen the rate of metabolic acidosis as the neonatal outcome criterion and the rate of operative deliveries and operative deliveries for fetal distress as maternal outcome criteria.

The Plymouth trial: a key phase 3 trial

In 1993, Westgate et al[59] in Plymouth, United Kingdom, published the first randomized controlled trial (RCT) comparing CTG plus ST analysis with CTG alone. In this trial 2434 high-risk women in labor were randomly allocated to CTG plus ST analysis (CTG plus ST) or to conventional CTG alone.

Women were eligible to enrollment in the trial if their gestational age was >34 weeks and there was an indication to apply a fetal scalp electrode for continuous internal cardiotocography (high-risk pregnancies including preeclampsia, antepartum hemorrhage, growth restriction, diabetes, previous cesarean delivery, abnormal antenatal or early labor external cardiotocogram, induced or augmented labor, presence of meconium, epidural analgesia, or breech presentation).

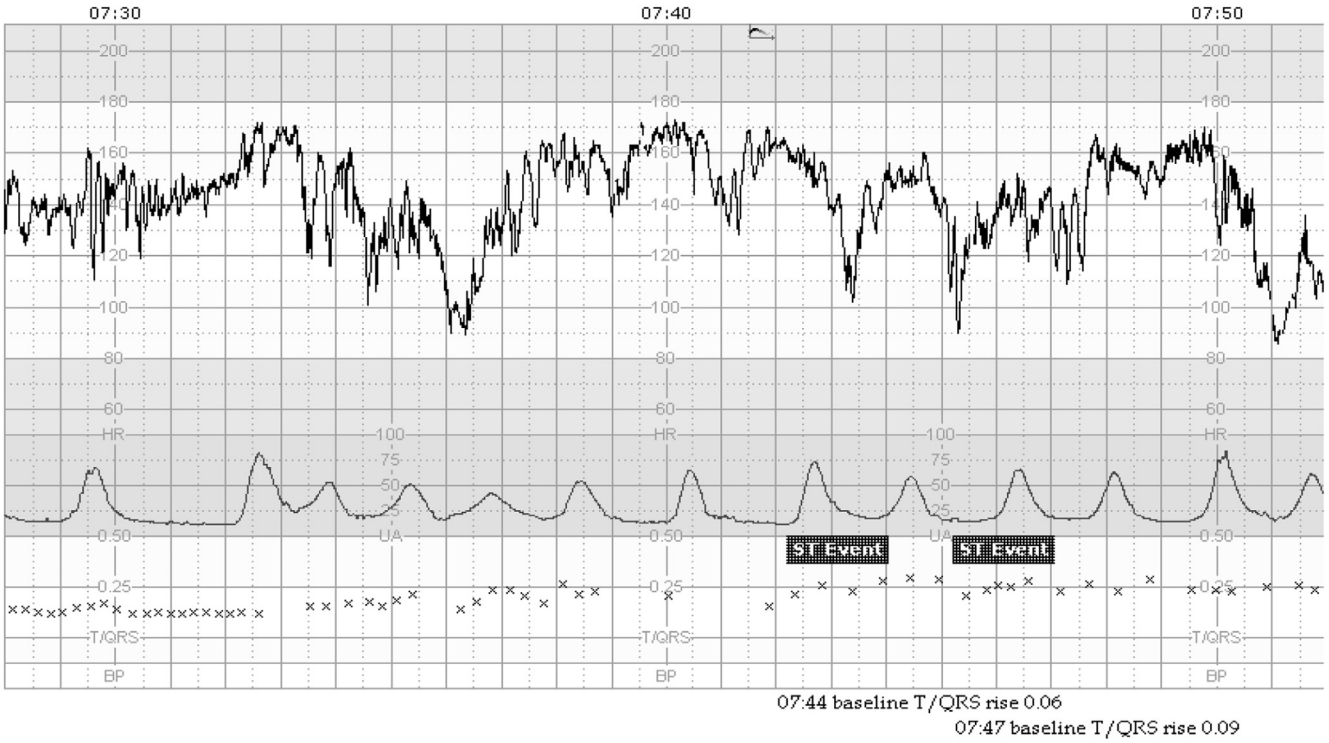
The primary outcome was the rate of operative delivery for fetal distress. Total operative delivery for fetal distress in the CTG plus ST arm was 5.0% (61 of 1219), a 49% fall, compared with 9.1% (111 of 1215) in the CTG-alone arm ($P < .001$; Table 1). The rate of metabolic acidosis was 1.1% in the CTG-alone arm (13 of 1215) and did not differ significantly with the CTG plus ST analysis arm (0.4%, 5 of 1219, odds ratio [OR], 2.63, 95% confidence interval [CI], 0.93−7.39, $P = .09$).

In this study an old analog signal processor was used, and the management guidelines were based on visual (and nonautomatic) identification of ST segment depression/trend analysis, and automatic T/QRS calculations. The ST event displayed on the screen with today's system was not available; consequently, the warning flags were lacking, requiring very close visual analysis of the changes occurring.

The Plymouth trial was of paramount importance for the development of signal processing and clinical management guidelines. Together with the observational studies presented in Appendix A, it stimulated both system development focusing on enhanced signal quality and the development of a multimedia training program including certification, a well-known need in the field of intrapartum monitoring.

Extensive research allowed the development of this product in collaboration with several academic centers at the end of the 1990s with support from the European Union, providing an opportunity to collect Europe-wide data for system

FIGURE 6
Example of STAN recording during active labor with baseline T/QRS events



Example of STAN recording, showing an abnormal CTG pattern with 2 significant ST events, baseline T/QRS rise of 0.06 and 0.09, respectively (paper speed: 1 cm/min).

CTG, cardiotocography; *STAN*, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

development.[77,78] The dissemination of knowledge about the pathophysiology of fetal adaptations to intrapartum hypoxia was considered essential because inadequate clinical response to CTG patterns[79] or to ST events would make the technology susceptible to the same pitfalls as CTG has presented, including failure to comply with guidelines.[80]

The European Union–financed project had the primary objective of disseminating knowledge about fetal physiological responses to hypoxia and ST segment analysis.[81] Ten European academic hospitals served as regional hubs of experience and centers of excellence. The interactive multimedia teaching and training package was called the STAN viewer program. The centers of excellence subsequently undertook their own projects for training and certification, and some conducted clinical studies and trials.

Four RCTs comparing CTG plus ST analysis with CTG alone (2000–2011)

Between 2000 and 2010, 4 RCTs were conducted in Europe (Sweden, Finland, France, and The Netherlands) with a later version of STAN comparing CTG plus ST with CTG alone among pregnant women in labor with an indication for continuous internal fetal monitoring (a selection of high-risk women).

The Swedish RCT[60] was conducted between December 1998 and June 2000 in 3 university hospital labor wards. Eligible participants were women in active labor at more than 36 weeks, with singleton fetuses in the cephalic presentation, and about whom a clinical decision has been made to apply fetal scalp electrode for continuous internal cardiotocography (women with suspicious or abnormal external cardiotocography, induced or oxytocin-augmented labor, meconium-stained amniotic fluid, or epidural analgesia).

A total of 4966 women enrolled in the trial. The main outcome measure was the rate of metabolic acidosis defined as pH <7.05 and base deficit >12 mmol/L, and the secondary outcome was the rate of operative delivery (cesarean delivery or instrumental delivery) for fetal distress. Prevalence of metabolic acidosis was 1.4%. Rates of metabolic acidosis was 0.7% (15 of 2159) in the CTG plus ST arm and 2% (31 of 2079) in the CTG arm (relative risk [RR], 0.47; 95% CI, 0.25–0.86; P=.02). Rates of operative delivery for fetal distress was 8.0% (193 of 2519) and 9% (227 of 2447), respectively (RR, 0.83; 95% CI, 0.69–0.99; P = .047) (Table 1). In the Appendix B, we present Interval between onset of FHR and ST changes and time of delivery, related to perinatal outcome and

TABLE 1
Overview of randomized clinical trials comparing CTG with CTG plus ST

Authors/year	Number of obstetric units, country, design	n	Primary outcome	Main results
Westgate et al, 1993[59]	1, United Kingdom, CTG plus ST analysis vs CTG only	2434	Total operative delivery rate, fetal distress and metabolic acidosis: umbilical artery (pH <7.05 and BD_{ecf} >12 mmol/L	Trend towards decrease in metabolic acidosis (CTG, 13/1215, CTG plus ST analysis, 5/1219, OR, 0.38; 95% CI, 0.13−1.07; $P = .09$) Decrease in total operative delivery for fetal distress rate (CTG, 111/1215, CTG plus ST analysis, 61/1219, OR, 0.51, 95% CI, 0.37−0.70; $P < .001$) No difference in rate of fetal blood sampling. (CTG, 114/1215, CTG plus ST analysis, 93/1219, OR, 0.80; 95% CI, 0.60−1.06; $P = .14$)
Amer-Wåhlin et al, 2001[60]	3, Sweden, CTG plus ST analysis vs CTG only	4966	Metabolic acidosis: umbilical artery (pH <7.05 and BD_{ecf} >12 mmol/L)	Decrease in metabolic acidosis (CTG, 31/2079, CTG plus ST analysis, 15/2159, OR, 0.46; 95% CI, 0.25−0.86; $P = .02$) Decrease in operative delivery for fetal distress rate (CTG, 227/2447, CTG plus ST analysis, 193/2519, RR, 0.83; 95% CI, 0.69−0.99; $P = .047$) No difference in rate of fetal blood sampling (CTG, 261/2079, CTG plus ST analysis, 234/2519, RR, 0.87; 95% CI, 0.74−1.03; $P=.12$)
Ojala et al, 2006[61]	1, Finland, CTG plus ST analysis vs CTG only	1483	Umbilical artery pH <7.10	No difference in metabolic acidosis. (CTG, 5/722, CTG plus ST analysis, 12/714, RR, 2.43; 95% CI, 0.86−6.85; $P = .093$) No difference in operative delivery rate (cesarean delivery: CTG, 35/739, CTG plus ST analysis, 47/733, RR, 1.35; 95% CI, 0.86−2.07; $P = .12$). Vacuum outlet rate: CTG, 79/739, CTG plus ST analysis, 70/733, RR, 0.89; 95% CI, 0.66−1.21, $P = .53$) Decrease in fetal blood sampling (CTG, 115/739, CTG plus ST analysis, 51/733, RR, 0.45; 95% CI, 0.33−0.61; $P < .001$)
Vayssière et al, 2007[62]	2, France, CTG plus ST analysis vs CTG only	799	Operative deliveries for fetal distress	No difference in operative deliveries (CTG, 221/400, CTG plus ST analysis 216/399, RR, 0.98; 95% CI, 0.86−1.11) Decrease in fetal blood sampling. (CTG, 248/400, CTG plus ST analysis, 108/399, RR, 0.44; 95% CI, 0.36−0.52)
Westerhuis et al, 2010[63]	9, The Netherlands, CTG plus ST analysis vs CTG only	5681	Metabolic acidosis in umbilical artery (pH <7.05 and BD_{ecf} >12 mmol/L)	No difference in metabolic acidosis. (CTG, 30/2840, CTG plus ST analysis, 20/2827, RR, 0.70; 95% CI, 0.38−1.28) No difference in operative delivery rate (CTG, 822/2840, CTG plus ST analysis, 789/2827, RR, 0.96; 95% CI, 0.87−1.06) No difference in low Apgar score or neonatal enceophalopatthy. 48% decrease in fetal blood sampling (CTG, 578/2840, CTG plus ST analysis, 301/2827, RR, 0.52; 95% CI, 0.46−0.59)
Belfort et al, 2015[70]	26, United States CTG plus ST analysis vs CTG only	11,108	Composite outcome (intrapartum fetal death, neonatal death, 5 minute Apgar score ≤3, umbilical artery pH <7.05 with BD_{blood} >12 mmol/L, intubation or ventilation at delivery, neonatal encephalopathy)	No difference in primary composite outcome (intrapartum fetal death, neonatal death, Apgar score <3 at 5 minutes, neonatal seizures, UA blood pH ≤7.05, with BD ≥12 mmol/L, intubation for ventilation or neonatal encephalopathy) (CTG, 40/5576, CTG plus ST analysis, 52/5532, RR, 1.31; 95% CI, 0.87−1.98; $P = .20$) No difference in operative delivery rate (CTG, 22.0%, CTG plus ST analysis, 22.8%, $P = .31$)

BD, base deficit; BD_{ecf}, BD in the extracellular fluid; *CI*, confidence interval; *CTG*, cardiotocography; *OR*, odds ratio; *RR*, risk ratio.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

metabolic acidosis in cases with adverse/complicated neonatal outcome in the swedich trial.[60]

Challenges to the trial's initial results led to reviews of both the database and the data analyses. In a publication discussing methodological challenges in evaluating medical technology in RCTs, a standard intention-to-treat (ITT) analysis was considered to be the most appropriate and was used to reanalyze the results.[82] In the standard ITT analysis, 83 allocated cases did not fulfill the inclusion criteria. These were cases of delivery before 36 weeks of gestation, breech delivery, malformations, and cases included after the trial was closed. Furthermore, the standard ITT analysis was extended to include not only validated umbilical cord blood samples but also all cord blood gas analyses, neonatal blood tests showing metabolic acidosis, and imputed data in cases with missing acid-base data.

The standard ITT analysis included all randomized women in the group in which they were randomly assigned. The data revision consisted of reclassifying all the cases and 12 misclassified cases were corrected.[82] However the reanalysis confirmed the initial significant results[60,82] in favor of ST plus CTG over CTG only.

Between 2003 and 2006, 2 more RCTs were conducted, one in Finland[61] in 2003−04 and the other in France during 2004−2006.[62] There were methodological differences between the French trial on the one hand and the Swedish and the Finnish trials on the other hand. In the Finnish RCT, women were eligible if they were in active labor, ≥36 weeks' gestation, had a singleton fetus in a cephalic presentation, and amniotomy had been decided. Exclusion criteria were contraindication for scalp electrode and admission to the labor ward in the second phase of labor.

The Finnish RCT (n = 1483) aimed to determine whether CTG plus ST analysis could reduce the rate of neonatal acidosis with a cord arterial pH <7.10, operative interventions (cesarean or instrumental deliveries), and need for fetal blood sampling. The prevalence of metabolic acidosis was 0.7% in the CTG arm, a lower-risk population compared with the control arm of the Swedish trial.[60] The French RCT enrolled women in labor with singleton fetus at gestational age ≥36 weeks in cephalic presentation with suspicious or abnormal CTG or meconium-stained amniotic fluid.

The aim of the French RCT was to determine whether CTG plus ST analysis reduced the rate of operative delivery for nonreassuring fetal status. A total of 799 patients were included. The prevalence of metabolic acidosis was 1.3% in the CTG-only arm. Neither the Finnish nor the French study had the power to detect a difference in neonatal acidemia nor operative interventions between the groups (Table 1). In both trials, a significant reduction in fetal blood sampling in the CTG plus ST analysis group was observed.

A large multicenter Dutch RCT was conducted between January 2006 and July 2008 to compare the incidence of metabolic acidosis (main outcome measure) between CTG plus ST and CTG-alone across 9 centers (n = 5681).[63] Women were eligible if in labor, aged ≥18 years with a singleton high-risk pregnancy (including hypertensive disorders, fetal growth restriction, premature rupture of membranes >24 hours, postdate pregnancy, failure to progress, need for pain relief, meconium-stained amniotic fluid, or nonreassuring CTG), had a fetus in cephalic presentation, and gestational age ≥36 weeks.

The prevalence of metabolic acidosis was 1.0% (30 of 2840) in the CTG arm. The Dutch RCT found lower but nonsignificant rate of metabolic acidosis in the CTG plus ST arm (0.7%, 20 of 2827) compared with the CTG-only arm (RR, 0.70 (range, 0.38−1.28)). In 2011, the authors published a revised version of their initial findings including corrected neonatal outcome numbers but the results remained largely unchanged.[83]

Table 1 provides an overview of the original publications of the 4 RCTs, together with the first one (the Plymouth trial[59]) and the last one (conducted in the United States[70]).

Conflicting results among the first 5 RCTs

Even without considering the methodological differences between the RCTs such as statistical power, definitions of criteria, eligibility criteria, or effect size estimation and the revised publications with corrected numbers, the first 5 RCTs still produced conflicting results:

- Of the 5 RCTs, only the Swedish RCT showed a significant reduction in metabolic acidosis in the CTG plus ST analysis group compared with the CTG-only group both in the initial publication[60] and the revised analysis.[82] However, the Finnish and the French RCTs were not designed or powered to detect significant differences in metabolic acidosis.[61,62]

- The rates of operative deliveries for fetal distress were significantly lower in the CTG plus ST arms in both the Plymouth and Swedish trials, whereas no differences were found between groups in the other RCTs.

Six meta-analyses (MAs) of the first five RCTs (2012−2014) (Table 2)

In general, clinicians and policy makers faced with conflicting study results on the same research question find it difficult to make decisions on interventions. Clinicians in particular cannot draw conclusion on the effectiveness of ST analysis because of the conflicting results of the five RCTs. Pooling results from 2 or more studies using a meta-analytic approach enables the synthesis of information collected through multiple studies into an overall estimation of the intervention effect.

Additional advantages of a meta-analysis include the increase in statistical power and consequently improvement in precision of estimates. By cumulating data, it is possible to detect effects that could not be detected in the individual studies because of a lack of power. However, conducting a meta-analysis requires a rigorous stepwise approach and collaboration between clinicians and statisticians. In the Supplemental Materials, we explained the critical aspects of the MA process and underline how different options may lead to varying or even divergent results.

TABLE 2
Meta-analyses published between 2012 and 2016 comparing CTG plus ST analysis with CTG alone

Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
European meta-analysis	Becker et al, 2013[66]	Type of studies: RCTs Population: women in labor with singleton fetus in cephalic presentation beyond 34 weeks of gestation	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as umbilical cord pH <7.05 and base deficit in the extracellular fluid >12 mmol/L, operative deliveries, Apgar score <7	Aggregate	RR, 0.72, 95% CI, 0.43–1.19, n = 15,338	RR 0.94, 95% CI 0.89–0.99, n = 15,338	Not performed	RR 0.86, 95% CI, 0.76–0.97 n = 15,338
US meta-analysis	Potti and Berghalla 2012[64]	Type of studies: RCTs Population: pregnant women with singleton fetus in cephalic presentation at term or near term in labor	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as umbilical arterial pH <7.05 and base deficit >12 mmol/L	Aggregate	RR, 0.80, 95% CI, 0.44–1.47, n = 14,574	Not performed	Not performed	Not performed
Swedish meta-analysis	Salmelin et al, 2013[65]	Type of studies: not precised. RCTs, observational and case-control studies included Population: women in active labor with term pregnancy, a singleton fetus in cephalic presentation, ruptured membranes, and the need for continuous fetal monitoring	Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as umbilical cord pH <7.05 and base deficit >12 mmol/L, hypoxic ischemic encephalopathy, cesarean delivery, instrumental vaginal delivery, fetal scalp blood sampling	Aggregate	RR 0.96, 95% CI, 0.49–1.88, n = 12,868	Not performed	RR, 0.93, 95% CI, 0.80–1.08, n = 12,904	Not performed
Cochrane meta-analysis	Neilson, 2012[67]	Type of studies: RCTs Population: pregnant women (and their fetuses) in labor, with a perceived need for continuous electronic fetal heart rate monitoring	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Cesarean delivery, metabolic acidosis defined as cord artery pH <7.05 and base deficit >12 mmol/L (blood), neonatal encephalopathy	Aggregate	RR, 0.78, 95% CI, 0.44–1.37, n = 14,574	Not performed	Not performed	Not performed

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

(continued)

TABLE 2
Meta-analyses published between 2012 and 2016 comparing CTG plus ST analysis with CTG alone *(continued)*

Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
IPD meta-analysis	Schuit et al, 2013[68]	Type of studies: phase III trials focusing on T/QRS ratio changes Population: women in labor with term singleton in cephalic presentation with an indication for internal electronic fetal monitoring	Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as an umbilical cordartery pH <7.05 and a base deficit calculated in the extracellular fluid >12 mmol/L	Individual patient data	RR, 0.76, 95% CI, 0.53–1.10, n = 12,987	RR, 0.94, 95% CI, 0.88–1.01, n =12987	RR, 0.94, 95% CI, 0.84–1.05, n = 12,987	RR, 0.91, 95% CI, 0.80–1.05, n = 12,987
Olofsson meta-analysis	Olofsson et al, 2014[84]	Type of studies: RCTs Population: not precised	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 Dutch RCT, 2010	Those events in which we found critical differences between the RCTs [see the accompanying part I review], and when improper handling of RCT data or errors were found in the 5 MAs	Aggregate	RR, 0.61, 95% CI, 0.41–1.07, n = 14,586	RR, 0.93, 95% CI, 0.88–0.99 (random-effect model), n = 14,539	Not performed[a]	Not performed[b]
Cochrane meta-analysis	Neilson, 2015[71]	Type of studies: RCTs Population: pregnant women (and their fetuses) in labor, with a perceived need for continuous electronic fetal heart rate monitoring	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010 US RCT, 2015	Cesarean delivery, cord artery pH <7.05 and base deficit >12 mmol/L, neonatal encephalopathy	Aggregate	RR, 0.72, 95% CI, 0.43–1.20, n = 25,682	Not performed	Not performed	Not performed

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019. *(continued)*

TABLE 2
Meta-analyses published between 2012 and 2016 comparing CTG plus ST analysis with CTG alone *(continued)*

Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
Sacconne meta-analysis	Saccone et al, 2016[72]	Type of studies: RCTs Population: singleton gestations in cephalic presentation at term or near term in labor	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010 US RCT, 2015	Perinatal composite outcome defined as at least 1 of the followings: intrapartum fetal death, neonatal death, Apgar score $\leq$3 at 5 minutes, neonatal seizure, metabolic acidosis (defined as umbilical arterial pH $\leq$7.05, and extracellular fluid base deficit $\geq$12 mmol/L), intubation for ventilation at delivery, or neonatal encephalopathy	Aggregate	First definition: RR, 0.74, 95% CI, 0.54–1.02,[c] n = 24,095 Second definition: RR, 0.81, 95% CI, 0.44–1.46,[d] n = 15,302	RR, 0.93, 95% CI, 0.86–1.01, n = 26,446	Not performed	Not performed
Blix meta-analysis	Blix et al, 2016[88]	Type of studies: RCTs Population: women in labor, $\geq$36 weeks of gestation with a singleton fetus in cephalic presentation and a decision for continuous electronic fetal monitoring during labor	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010 US RCT, 2015	Operative delivery for fetal distress, metabolic acidosis defined as pH $<$7.05 and base deficit in extracellular fluid $>$12 mmol/L	Aggregate	OR, 0.64, 95% CI, 0.46–0.88, n = 26,493	RR, 0.96, 95% CI, 0.91–1.02, n = 26,446	RR, 0.88, 95% CI, 0.75–1.03, n = 26,446	RR, 0.87, 95% CI, 0.74–1.03, n = 26,446
Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
European meta-analysis	Becker et al, 2012[66]	Type of studies: RCTs Population: women in labor with a singleton fetus in cephalic presentation beyond 34 weeks of gestation	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as umbilical cord pH $<$7.05 and base deficit in the extracellular fluid $>$12 mmol/L, operative deliveries, Apgar score $<$7	Aggregate	RR, 0.72, 95% CI, 0.43–1.19, n = 15,338	RR, 0.94, 95% CI, 0.89–0.99, n = 15,338	Not performed	RR, 0.86, 95% CI, 0.76–0.97, n = 15,338

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

(continued)

TABLE 2
Meta-analyses published between 2012 and 2016 comparing CTG plus ST analysis with CTG alone *(continued)*

Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
US meta-analysis	Potti and Berghalla 2012[64]	Type of studies: RCTs Population: pregnant women with singleton fetus in cephalic presentation at term or near term in labor	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as umbilical arterial pH <7.05 and base deficit >12 mmol/L	Aggregate	RR 0.80, 95% CI 0.44–1.47, n = 14,574	Not performed	Not performed	Not performed
Swedish meta-analysis	Salmelin et al, 2013[65]	Type of studies: Not precise. RCTs, observational and case-control studies included Population: women in active labor with term pregnancy, a singleton fetus in cephalic presentation, ruptured membranes, and the need for continuous fetal monitoring	Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as umbilical cord pH <7.05 and base deficit >12 mmol/L, hypoxic ischemic encephalopathy, cesarean delivery, instrumental vaginal delivery, fetal scalp blood sampling	Aggregate	RR, 0.96, 95% CI, 0.49–1.88, n = 12,868	Not performed	RR, 0.93, 95% CI, 0.80–1.08, n = 12904	Not performed
Cochrane meta-analysis	Neilson, 2013[67]	Type of studies: RCTs Population: pregnant women (and their fetuses) in labor, with a perceived need for continuous electronic fetal heart rate monitoring	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Cesarean delivery, metabolic acidosis defined as cord artery pH <7.05 and base deficit >12 mmol/L (blood), neonatal encephalopathy	Aggregate	RR, 0.78, 95% CI, 0.44–1.37, n = 14,574	Not performed	Not performed	Not performed

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019. *(continued)*

TABLE 2
Meta-analyses published between 2012 and 2016 comparing CTG plus ST analysis with CTG alone *(continued)*

Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
IPD meta-analysis	Schuit et al, 2013[68]	Type of studies: phase III trials focusing on T/QRS ratio changes Population: women in labor with a term singleton fetus in cephalic presentation with an indication for internal electronic fetal monitoring	Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010	Metabolic acidosis defined as an umbilical cordartery pH <7.05 and a base deficit calculated in the extracellular fluid >12 mmol/L	Individual patient data	RR, 0.76, 95% CI, 0.53–1.10, n = 12,987	RR, 0.94, 95% CI, 0.88–1.01, n = 12,987	RR, 0.94, 95% CI, 0.84–1.05, n = 12,987	RR, 0.91, 95% CI, 0.80–1.05, n = 12,987
Olofsson meta-analysis	Olofsson et al, 2014[69,84]	Type of studies: RCTs Population: not precised	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 Dutch RCT, 2010	"Those events where we found critical differences between the RCTs [see the accompanying part I review], and when improper handling of RCT data or errors were found in the 5 MAs"	Aggregate	RR, 0.61, 95% CI, 0.41–1.07, n = 14,586	RR, 0.93, 95% CI, 0.88–0.99 (random-effect model), n = 14,539	Not performed[a]	Not performed[b]
Cochrane meta-analysis	Neilson, 2015[71]	Type of studies: RCTs Population: pregnant women (and their fetuses) in labor, with a perceived need for continuous electronic fetal heart rate monitoring	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010 US RCT, 2015	Cesarean delivery, cord artery pH <7.05 and base deficit >12 mmol/L, neonatal encephalopathy	Aggregate	RR, 0.72, 95% CI, 0.43–1.20, n = 25,682	Not performed	Not performed	Not performed

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019. *(continued)*

TABLE 2
Meta-analyses published between 2012 and 2016 comparing CTG plus ST analysis with CTG alone *(continued)*

Designation	Authors/year	Eligibility criteria	Studies considered in the quantitative synthesis	Primary outcome	Type of meta-analysis	Metabolic acidosis	Total operative delivery	Total operative delivery for fetal distress	Operative vaginal delivery for fetal distress
Sacconne meta-analysis	Saccone et al, 2016[72]	Type of studies: RCTs Population: singleton gestations in cephalic presentation at term or near term in labor	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010 US RCT, 2015	Perinatal composite outcome defined as at least 1 of the following: intrapartum fetal death, neonatal death, Apgar score $\leq$3 at 5 minutes, neonatal seizure, metabolic acidosis (defined as umbilical arterial pH $\leq$7.05 and extracellular fluid base deficit $\geq$12 mmol/L), intubation for ventilation at delivery, or neonatal encephalopathy	Aggregate	First definition: RR, 0.74, 95% CI, 0.54–1.02,[c] n = 24,095 Second definition: RR, 0.81, 95% CI, 0.44–1.46,[d] n = 15,302	RR, 0.93, 95% CI, 0.86–1.01, n = 26,446	Not performed	Not performed
Blix meta-analysis	Blix et al, 2016[73]	Type of studies: RCTs Population: women in labor, $\geq$36 weeks of gestation with a singleton fetus in cephalic presentation and a decision for continuous electronic fetal monitoring during labor	Plymouth RCT, 1993 Swedish RCT, 2001 Finnish RCT, 2006 French RCT, 2007 Dutch RCT, 2010 US RCT, 2015	Operative delivery for fetal distress, metabolic acidosis defined as pH $<$7.05 and base deficit in extracellular fluid $>$12 mmol/L	Aggregate	OR, 0.64, 95% CI, 0.46–0.88, n = 26,493	RR, 0.96, 95% CI, 0.91–1.02, n = 26,446	RR, 0.88, 95% CI, 0.75–1.03, n = 26,446	RR, 0.87, 95% CI, 0.74–1.03, n = 26,446

CI, confidence interval; *CTG*, cardiotocography; *IPF*, individual patient data; *n*, number of patients in the analysis; *MA*, meta-analysis; *OR*, odds ratio; *RCT*, randomized controlled trial; *RR*, relative risk;

[a] But fetal distress among all operative deliveries was analyzed; [b] But fetal distress among all operative vaginal deliveries was analyzed; [c] Neonatal metabolic acidosis was defined as umbilical arterial pH $\leq$7.05 and extracellular fluid base deficit $\geq$12 mmol/L (Plymouth trial excluded); [d] Neonatal metabolic acidosis was defined as umbilical arterial pH $<$7.05 and extracellular fluid base deficit 12 mmol/L (US trial excluded).

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Between 2012 and 2013, 5 MAs comparing the effect of CTG plus ST analysis vs CTG alone on both neonatal and maternal outcomes were published. These meta-analyses combine data collected from trials including high-risk women in labor with a term singleton pregnancy in cephalic presentation. Four MAs were performed with the use of aggregated data (with authors from the United States,[64] Sweden,[65] Europe,[66] and the Cochrane Collaboration[67]), while 1 was an individual patient data (IPD) MA.[68] Three aggregate MAs focused on RCTs evaluating the effectiveness of the ST analysis,[64,66,67] while 1 MA included RCTs as well as observational studies and case-control studies.[65] The IPD MA[68] selected phase III trials. Furthermore, whereas two MAs (the US MA[64] and the IPD MA[68]) have chosen the rate of metabolic acidosis as the primary outcome criteria of their analysis, the other MAs cumulated several outcome criteria.

No differences between groups were found in the rate of metabolic acidosis among the 5 MAs. Four MAs (the European,[66] the American,[64] the Cochrane,[67] and the IPD MA[68]) compared the rate of operative vaginal delivery between the randomization groups. All found a significant reduction in the rate of operative vaginal delivery (from 10% in the IPD MA[68] to 12% in the European MA[66]). There were inconsistencies among the MAs in the effect of CTG plus ST on operative vaginal delivery for fetal distress (with a significant reduction in the European MA[66] but not in the IPD MA[68]).

Because of its approach based on individual data rather than aggregated data, the IPD MA of Schuit et al[68] was expected to present several advantages over the 3 aggregate MAs. Indeed, the IPD MA enabled the authors to study outcomes, which were not reported in any of the 5 RCTs, to perform subgroup analyses and to impute the missing data, thus limiting the potential bias because of missing data. Nonetheless, Schuit's IPD MA has some limitations reported in 2014 by an international collaboration, Olofsson et al.[69]

In this sixth MA of the first 5 RCTs that were published at the time, Olofsson et al[69] systematically appraised the quality, strengths, and weaknesses of the individual RCTs in a first publication.[84] In the second publication, the authors appraised the methodology, quality, and conduct of the MAs published so far and explained the sources of their conflicting results.[69]

In this aggregate MA, critical differences between RCTs were found, and an improper handling of RCT data and errors were found.[69] Differences among the 5 MAs leading to conflicting results were noted in 3 stages of the process of meta-analysis: study selection, data extraction, and statistical analysis. The differences observed between the 5 MAs, which Olofsson et al[69] has listed are summarized in the following text and grouped according to the stage of the MA process to which they correspond.

First, criteria for the inclusion of RCTs varied. Three MAs included all 5 RCTs,[64,66,67] while 2 excluded the Plymouth trial.[59] The IPD MA[68] and the Swedish MA[65] excluded the Plymouth trial[59] because the automatic computerized ST event was not available. Moreover, decisions to exclude some RCTs for specific outcomes differed among the 5 MAs: for example, the Cochrane meta-analysis excluded the Dutch RCT[63] when analyzing fetal blood sampling and operative vaginal delivery outcomes,[67] but no explanation was given for this exclusion.

Second, Olofsson et al[69] pointed out the inconsistencies in the numbers extracted from the individual RCTs between the MAs. For the purpose of the present review, we focus our discussion on 2 of these inconsistencies only, namely metabolic acidosis and operative deliveries. Three MAs, Cochrane,[67] the American,[64] and European[66] MAs, did not use the revised ITT analysis of the Swedish RCT with respect to metabolic acidosis.[82]

In the European meta-analysis,[66] contrary to what was described in the text, the figures related to the Swedish RCT did not correspond to the revised ITT analysis; the Cochrane MA[80] cited the revised version of the Swedish RCT[82] but used the initial Swedish data in their analysis,[60] and the US MA[64] cited and used the initial data of the Swedish trial.[60]

There were errors in the numbers extracted from the primary RCTs and used in the MAs. For example, the pooled numbers of cesarean and operative vaginal deliveries from the Swedish, Finnish, French, and Dutch RCTs presented in the IPD MA are erroneous and do not tally.[68]

There is a further error in the number of cases of metabolic acidosis extracted from the French trial[62] and presented in the IPD MA.[68] The French trial reported 13 cases of metabolic acidosis (8 in the CTG plus ST group vs 5 cases in the CTG-only group), whereas the IPD included 19 cases. The definition of metabolic acidosis differed in the individual trials and resulted in inconsistencies across MAs.

In particular, 2 algorithms are available for the calculation of BD and were used by the primary trials. They are BD in blood (BD_{blood}) or BD in the extracellular fluid (BD_{ecf}) and produce different results, with the BD_{blood} giving a higher value by 2−3 mEq/L. The Cochrane MA[67] defined metabolic acidosis as a pH <7.05 and a BD_{blood} >12 mmol/L in the umbilical cord artery, but the figures presented in the tables correspond to a mixture of metabolic acidosis based on BD_{blood} for the Finnish and the Dutch RCTs and based on BD_{ecf} for the Plymouth, Swedish, and French RCTs; the US[64] and the Swedish MAs[65] used mixed measures of metabolic acidosis based on BD_{ecf} and those based on BD_{blood} for the Finnish RCT.

Finally, strategies for handling missing data were heterogeneous, particularly with respect to metabolic acidosis[69,84]: the Swedish[65] and US[64] MAs were based on published data from the 5 RCTs, while for the other MAs, missing data could be provided or imputed (for the IPD MA[68]).

Third, the way statistical analysis was conducted varied substantially. As presented in the Supplementary Material, choosing the estimation method for the pooled treatment effect (here the effect of CTG plus ST analysis compared with CTG alone) is based on statistical

considerations. Each statistical model used to estimate the pooled treatment effect is based on several hypotheses that have to be validated.

One issue concerns the statistical heterogeneity between studies (ie, the variability of the treatment effect across studies). If considered substantial, this variability should be taken into account by using random-effect models. If not, the statistical validity of the pooled estimate of treatment effect will be biased. Although the 5 MAs (including the IPD MA[68]) assessed statistical heterogeneity of the estimates of the ST analysis effectiveness among studies, only 3[64,66,67] used this estimation appropriately to choose between random- and fixed-effects models for data synthesis, and even these 3 used different cutoff values to choose among these model types.

Olofsson et al[69] corrected these data-handling and methodological errors and conducted a further meta-analysis of the 5 RCTs, which showed that compared with CTG alone, CTG plus ST significantly reduced the rate of metabolic acidosis by 39% (RR, 0.61; 95% CI, 0.41−0.91; $P = .02$) and operative delivery rate by 7% (RR, 0.93; 95% CI, 0.88−0.99; $P = .02$).

The US trial (2015)

With 11,108 participants enrolled in 26 centers, the US RCT was the largest to compare CTG plus ST with CTG alone (Table 1).[70] Women were eligible to participate if they were in labor with a singleton, cephalic fetus, at ≥36 weeks' gestational age. Thus, contrary to the trials published so far, the US trial did not focus on high-risk pregnancies requiring continuous internal fetal monitoring. The primary outcome was a composite outcome including intrapartum fetal death, neonatal death, Apgar score of ≤3 at 5 minutes, neonatal seizure, an umbilical artery blood pH of ≤7.05 with a base deficit of ≥12 mmol/L, intubation for ventilation at delivery, or neonatal encephalopathy.

No differences were found between groups in the composite outcome. Likewise, there were no differences in cesarean or operative delivery rates. Metabolic acidosis occurred in 0.05% (3 of 5532) of the cases in the CTG plus ST arm and 0.14% (8 of 5576) in the CTG-only arm, but the difference was not statistically significant (RR, 0.37; 95% CI, 0.10−1.41; $P = .13$). This trial is difficult to compare with those conducted in Europe because of substantial methodological differences and a major change in guidelines:

1. The conduct of the trial across 26 centers took more than 3.5 years (41 months). The very low inclusion rate in some centers suggests insufficient experience with the CTG plus ST technology.
2. Recruitment was limited to weekdays between 8:00 AM and 2:00 PM, and the population was mainly women at low risk for fetal hypoxia. With 43,376 patients assessed for eligibility and 32,268 (74.4%) not enrolled, selection bias is a substantial possibility.
3. The composite outcome was unexpectedly low in both arms of the study: 0.94% in the CTG plus ST arm and 0.72% in the CTG arm, compared with the expected rate of 1.75%. This suggests a Hawthorne effect.[85]
4. The clinical guidelines for management of CTG plus ST events used in Europe were changed to a simplified version for the US trial, resulting in a lowering of the threshold for intervention. Specifically, it set a shorter duration (60 rather than 90 minutes) and lower magnitudes of ST changes for intervention. Moreover, the instructions to trialists stated, "Do not rely solely on the appearance of an ST event marker to signal the need for obstetrical intervention. If you suspect, on the basis of fetal heart rate only and/or clinical data that the fetus is experiencing severe hypoxia, you should manage the patient accordingly despite the absence of an ST event marker."

These instructions had several potential consequences. First, they were likely to have resulted in unnecessary interventions and may have affected the results. More importantly, it produced a study design that necessarily prevented the trial from answering the primary research question: does ST analysis of the fetal electrocardiogram, as an adjunct to conventional electronic fetal heart rate monitoring, reduce the risk of fetal compromise?

The substantially larger discretion for intervention not only kept the study from meeting its aim, but it also defeated the purpose of STAN use: to increase the diagnostic accuracy of intrapartum surveillance.

Finally, it resulted in magnifying all of the problems observed in the first half of many of the other RCTs and observational studies, described in the previous text, and simultaneously prevented any training effect. Essentially, instead of measuring the effect of STAN as intended to be used, it measured education about fetal physiology, which was the same in both groups.

Meta-analyses including the US trial and the resulting controversies (2015−2017) (Table 2)

After the publication of 5 RCTs between 1993 and 2010 and 6 MAs between 2012 and 2014, the negative results of the US trial published in August 2015 revived the debate on the benefits of STAN. Three further MAs of the now 6 available RCTs were published (Table 2). A Cochrane aggregate meta-analysis[71] was published in December 2015; an aggregate MA by Saccone et al[72] and an aggregate MA by Blix et al[73] complete with a trial sequential MA were published in January 2016. A trial sequential MA is a post hoc statistical method used to quantify the possibility of false-positive and false-negative results in a meta-analysis and to evaluate the reliability of findings from the pooled cumulative data of several trials.[86]

The main comparison of the Cochrane meta-analysis focused on CTG plus ST analysis vs CTG alone. RCTs that included women in labor, with a perceived need for continuous electronic fetal heart rate monitoring were selected. Primary endpoints were caesarean delivery, metabolic acidosis, and neonatal encephalopathy.

Saccone et al MA[72] compared the efficacy of CTG plus ST analysis with CTG

TABLE 3
Operative delivery rates for fetal distress in the Swedish trial separation of data according to mode of monitoring and phase of the trial[60]

Intention to treat	CTG plus ST	Control	OR, 95% CI, *P* value
First phase	114 8.6%	104 8.3%	1.03 n.s
Cases, n	1333	1250	
Second phase	79 6.7%	123 10.3%	0.62 0.46−0.85 .002
Cases, n	1186	1197	

CI, confidence interval; *CTG*, cardiotocography; *OR*, odds ratio.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

alone by including RCTs, which enrolled women with singleton, cephalic babies, in labor at or near term. The primary outcome was a perinatal composite outcome, defined as at least 1 of the following: intrapartum fetal death, neonatal death, Apgar score ≤3 at 5 minutes, neonatal seizure, metabolic acidosis (defined as umbilical arterial pH ≤7.05, and extracellular fluid base deficit ≥12 mmol/L), intubation for ventilation at delivery, or neonatal encephalopathy.

The meta-analysis by Blix et al[73] aimed to quantify the efficacy of CTG plus ST analysis compared with CTG alone. RCTs including women in labor, ≥ 36 weeks of gestation with a singleton fetus in a cephalic presentation, and a decision for continuous electronic fetal monitoring during labor were selected for analysis. Two endpoints were defined: metabolic acidosis and operative deliveries for fetal distress. Blix et al[73] have included all 6 RCTs and have conducted a sensitivity analysis in which they have excluded the Plymouth trial that used an old STAN technology and included women from 34 weeks' gestation[59] and the US trial that used a different algorithm for interventions.[70]

All MAs showed significant reductions in operative vaginal delivery rates (8% in the Blix et al[73] MA and in the Cochrane MA to 9% in the Saccone et al[72] MA). Blix et al[73] showed a significant reduction in metabolic acidosis (36%, OR, 0.64, 95% CI, 0.46−0.88), while the Cochrane[71] and the Saccone et al[72] reported nonsignificant reductions: the Cochrane MA showed a reduction of 28% (RR, 0.72, 95% CI, 0.43−1.20). Saccone et al used 2 definitions of metabolic acidosis: (1) pH ≤7.05 and BD_{ecf} ≥12 mmol/L, and (2) pH <7.05 and BD_{ecf} >12 mmol/L. The Plymouth RCT[59] was excluded from the analysis of metabolic acidosis analysis using definition 1 while the US RCT was excluded from the analysis of metabolic acidosis with definition 2. Whatever the definition used, nonsignificant reductions were reported (26% with definition 1, 19% with definition 2).

Why did these variations occur? First, they used slightly different data sets when analyzing the endpoint operative vaginal delivery. Saccone et al[72] wrongly used the number of operative vaginal delivery for nonreassuring fetal state from the French RCT,[74] whereas Blix et al[73] and the Cochrane MAs[71] correctly used operative vaginal deliveries for all indications; Saccone et al[72] excluded the US trial data from their analysis on metabolic acidosis because the US trial defined metabolic acidosis as a pH ≤7.05 and base deficit ≥12 mmol/L, whereas the 5 other trials defined metabolic acidosis as a pH <7.05 and base deficit >12 mmol/L.

We assume that only Blix et al[73] used the correct data. Specifically, revised data for the Swedish and Dutch trials were published years later,[82,83] and only Blix et al[73] used these revised data for MA on STAN monitoring. Moreover the Finnish trial,[61] unlike the others, initially presented results on metabolic acidosis with base deficit calculation in blood rather than in the extracellular fluid. The corrected data, provided elsewhere,[87] are required for comparability and were used by Blix et al.[73] Second, statistical methods differed in 3 key aspects: the effect measure (risk ratio vs odds ratio), the weighting of individual studies, and the recognition of heterogeneity among the studies.

Not surprisingly, the divergent results of the 3 meta-analyses prompted a heated debate, including 5 letters to the editor published in 2016.[88-93] Most of

TABLE 4
Inclusion rate of the 6 RCTs comparing STAN plus CTG vs CTG, only in labor

Study	Population risk status %[a]	Patients enrolled	Recordings per week	Metabolic acidosis rate in STAN arm, %
Plymouth (United Kingdom)	1.1	2434	31	0.4
Swedish	1.4	4966	21	0.6
Finland	0.6	1436	14	0.8
The Netherlands	1.0	5681	5	0.7
France	1.3	799	3	2.0
United States	0.14	11,108	1.2	0.05

CTG, cardiotocography; *RCT*, randomized controlled trial; *STAN*, ST segment analysis.

[a] Prevalence of metabolic acidosis in the control arm.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

TABLE 5
Observational studies reporting the clinical use of cardiotocography and ST analysis of fetal ECG in labor after the European Union program

Reference	Type of study	Number of obstetric units/country	n	Main results and conclusions
Kwee et al, 2004[102]	Prospective	1, The Netherlands	637	ST changes in all 5 cases with severe metabolic acidosis and in 46% of cases with mild metabolic acidosis; CTG plus ST more specific than CTG alone
Ross et al, 2004[103]	Retrospective	7 clinicians, United States	51	ST analysis in addition to CTG monitoring improves consistency in decision and timing of operative deliveries
Devoe et al, 2006[132]	Prospective	6, United States	530	The negative predictive value for avoiding operative deliveries in fetuses with a nonreassuring CTG pattern was 95%
Norén et al, 2006[105]	Prospective	2, Sweden	4830	Decrease in metabolic acidosis ($P < .05$)
Welin et al, 2007[87]	Retrospective	1, Sweden	1875	Decrease in overall and emergency cesarean delivery rates (OR, 0.18; 95% CI, 0.07–0.49). Low metabolic acidosis rate (0.5%)
Vayssière et al, 2007[133]	Prospective	2, France	411	The negative predictive value of predicting pH <7.05 in fetuses with a nonreassuring CTG pattern was 98%, sensitivity was 63%
Kale et al, 2008[106]	Prospective with historical controls	1, Singapore	474	Decrease in instrumental deliveries for fetal distress ($P < .025$), admissions to NICU ($P < .01$), and in low Apgar scores ($P < .001$). No difference in rate of cesarean delivery for fetal distress
Massoud et al, 2007[107]	Prospective	1, France	1889	Metabolic acidosis rate low (0.38%); trend to decrease in operative delivery for fetal distress
Doria et al, 2007[108]	Prospective	1, United Kingdom	1502	Metabolic acidosis occurred in 2.8%, 70% of which were identified by STAN. Retrospective analysis of the 14 cases with neonatal encephalopathy and monitored by STAN showed that the guidelines for intervention were not followed.
Palmgren ColovNS, 2007	Prospective	1, Denmark	1168	Number of operative deliveries decreased without an increase in the number of newborns with metabolic acidosis
Melin et al, 2008[110]	Retrospective case control	1, Sweden	506	An ST event occurred in 79% of severe and 75% of moderate metabolic acidemia cases and in 50% of controls. 2 of 3 cases with severe and <50% of cases with moderate metabolic acidemia were preceded by ST events coinciding with CTG abnormalities.
Rzepka et al, 2010[111]	Prospective	1, Poland	83	ST changes reflect metabolic acidosis better than CTG
Norén and Carlsson, 2010[112]	Prospective	1, Sweden	12,832	Metabolic acidosis decreased from 0.72% to 0.06%; unchanged rate of cesarean deliveries for fetal distress, decrease in overall cesarean delivery rate
Ragupathy et al, 2010[113]	Prospective	1, United Kingdom	253	1.5% rate of metabolic acidosis at birth, 100% preceded by ST events, but guidelines not followed. High percentage of ST events in laboring women
Doret et al, 2011[134]	Prospective	1, France	3112	No fetal blood sampling performed. 14 cases with metabolic acidosis, 11 of which were not managed according to guidelines. Cesarean delivery rate for fetal distress 9.5%. Study supports use of ST analysis without fetal blood sampling, but warns about guideline violations.
Kessler et al, 2013[135]	Prospective	1, Norway	6010	During the study period (4 y) metabolic acidosis rate decreased from 1.4% to 0.3% and the cesarean delivery rate from 10.1% to 8.8%.
Timonen et al[115]	Prospective	1. Finland	42,146	When the first 2 years of ST analysis usage were compared in the subsequent 9 years, there was a significant reduction in the rates of metabolic acidosis ($P = .0151$), cesarean deliveries ($P < .0001$), and fetal scalp blood sampling ($P < .0001$).
Landman et al[114]	Retrospective	1. The Netherlands	19,664	During the study period (14 years), the incidence of umbilical cord metabolic acidosis decreased from 2.5% to 0.4 % ($P < .001$). The fetal blood sampling rate and the total number of cesarean and instrumental vaginal deliveries ($P < .001$ for both)

CI, confidence interval; CTG, cardiotocography; NICU, neonatal intensive care unit; OR, odds ratio; RCT, randomized controlled trial; STAN, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

them recognized the Blix et al[73] MA to be the only one based on the accurate revised numbers.[91,92] Vayssiere et al[92] showed that use of CTG plus ST consistently reduced metabolic acidosis rates when the correct numbers from the six RCTs are used, regardless of the statistical analysis applied.

Because the meta-analysis by Blix et al[73] was the only one to use the correct numbers, as noted above, and also used the most appropriate methodology and because as Vayssiere et al[92] pointed out, using the accurate numbers produces significant findings for metabolic acidosis with all statistical methods, it must be concluded that CTG plus ST monitoring has important benefits over CTG alone.

Blix et al[73] reported that CTG plus ST reduced operative vaginal delivery rates by 8% (RR, 0.92; 95% CI, 0.86−0.99) (2862 of 26,446 events) and metabolic acidosis rates by 36% (OR, 0.64, 95% CI, 0.46−0.88) (151 of 26,446 events). That contrasts with the lack of benefit shown for other intrapartum fetal monitoring technologies (see the previous text). Accordingly, good evidence supports the continued use of ST analysis in intrapartum fetal monitoring.

Lessons learned and the way forward

Several important lessons can be learned from the experience of implementing ST analysis in practice and from the RCTs evaluating its efficacy. Some of these are considered below.

Diagnostic tools and their impact on outcomes

For diagnostic tools such as ST analysis, adequate use of the technology and strict adherence to the management guidelines are essential for an effect to be seen on outcomes. Good signal quality, adherence to interpretation guidelines, optimum decision making, and prompt action all have an impact on results. In the Swedish RCT, there were recommendations to start recording during the first stage of labor and to maintain recording until delivery or within 20 minutes of delivery. Reduction in metabolic acidosis exceeded 50% in cases adhering to these recommendations and increased over time as familiarity with the new method increased; both adequate signal quality and interpretation presented challenges to clinicians during the trial, especially during the first half, but these lessened thereafter. The rates of operative deliveries for fetal distress decreased significantly in the CTG plus ST arm after the interim analysis.

The larger reduction in metabolic acidosis in the CTG plus ST arm as well as the decreasing number of infants admitted to neonatal intensive care during the second period also underlines the importance of training.[94] Table 3 also shows the operative delivery rates for fetal distress during the different phases of the trial. Furthermore, the decreasing number of fetal blood samplings over time also suggests a gradual improvement in performance over the course of the study.

The Dublin trial of CTG monitoring showed a similar training effect, with results improving with experience.[95] During the second phase of the trial, the metabolic acidosis rate declined from 1.48% to 0.50% (OR, 0.33, $P = .045$). The IPD meta-analysis of CTG plus ST also showed significant reductions in metabolic acidosis, low cord artery pH, admission for neonatal intensive care, and composite adverse outcomes during the second half of the RCTs, reductions not observed in analyses of the initial study periods.[68,96]

Introducing a new technology such as ST analysis into clinical use requires rigorous training and sometimes a paradigm change regarding interpretation of information and clinical management. The learning loop approach, including regular case reviews, might usefully be tested in this context. Frequent case review was implemented as a learning strategy in the Swedish RCT and made it possible to detect and correct errors but also to solve the technical challenges that occurred. New guidelines sometimes require that existing clinical practices be challenged and new patterns of interaction instituted among healthcare professionals.

In trials comparing CTG plus ST to CTG only, the additional information provided by ST data may have caused marked changes in practices for the management of CTG patterns, which can take time to be embedded. Perception of liability is strong in modern obstetrics and may discourage innovation and adherence to protocols. Clinicians are known to differ in their approach to new technology, depending on age, sex, and other personal characteristics.[94]

The current analysis has focused on ST analysis. However, CTG is the predominant marker of adverse events. New CTG classification systems have been developed recently, raising the question whether CTG plus ST interpretation can still be valid if based on the new classification system.[12] At the introduction of the CTG plus ST analysis,[52] CTG patterns had to be used, regardless of their specific pathophysiological significance. Today, with access to large databases, there is an opportunity to identify CTG patterns associated with ST events and adverse outcome not covered by existing CTG standards. This indicates that CTG plus ST analysis may be undertaken, regardless of the CTG classification system used, as long as a more physiologically oriented approach is used for CTG assessment in connection with an ST event.[97]

Limitations of RCT design: enrollment rate and the Hawthorne effect

Trial enrollment rates matter for practitioners' acquisition of experience with the new technology and their becoming comfortable enough to change practice. The inclusion in multicenter studies of smaller sites with slow enrollment rates raises questions whether personnel at those sites acquired sufficient experience to use the technology to its full potential. Table 4 illustrates the impact of enrollment rates on clinical outcomes in the ST analysis trials.

Changes in individuals' behavior as a result of the awareness that they are being observed have been labeled the Hawthorne effect.[85] It is likely that this effect has had a strong impact on the RCTs evaluating intrapartum fetal monitoring. There is a strong focus on the capacity to give continuous attention to the tracing, and outcome is very

FIGURE 7
STAN clinical guidelines that prompt clinical intervention

ST-changes that prompt clinical intervention such as delivery or solving a cause of fetal distress[133]

	Intermediary CTG	Abnormal CTG
Episodic T/QRS-rise (duration < 10 min)	Increase > 0.15 from baseline	Increase > 0.10 from baseline
Baseline T/QRS-rise (duration ≥10 min)	Increase > 0.10 from baseline	Increase > 0.05 from baseline
Biphasic ST (a component of the ST-segment below the baseline)	Continuous >5 min or >2 episodes of coupled Biphasic ST type 2 or 3	Continuous >2 min or >1 episode of coupled Biphasic ST type 2 or 3

The ST log requires 20 minutes of recording for automatic ST analysis to start. A decrease in signal quality with insufficient number of T/QRS measurements requires manual data analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

dependent on how the situation is managed. A particular characteristic of the Hawthorne effect is that it tends to increase over time until it plateaus and then declines.[98,99]

Results of recent observational studies

The shortcomings in the generalizability of RCT results mentioned in the previous text[100] have led to increased attention to publications evaluating the use of ST analysis in the real-world settings.[101]

Several observational studies evaluating about 45,000 cases monitored with ST analysis in routine obstetric care have been published in recent years, and their results are summarized in Table 5.[62,87,102-113] Two additional large observational cohorts of women monitored with ST analysis were published, one in 2018 and one in 2019. In The Netherlands, data collected between 2000 and 2013 in a population of 19,664 high-risk women in labor showed a reduction of the metabolic acidosis rate from 2.5% in the first years to 0.4% in more recent years ($P < .001$) and a decrease in the total number instrumental vaginal deliveries from 13.8% to 10.3% ($P < .001$).[114] In Finland, data collected between 2001 and 2011 from 42,146 laboring women showed a 67% reduction in the incidence of metabolic acidosis between the first 2 years and the next 9 years in the ST analysis group. Of the 27 newborns with metabolic acidosis, 10 occurred during the initial 2 years (1.0% vs 0.25%, RR, 0.33; 95% CI, 0.15−0.72). Cesarean deliveries also decreased by 18% (17.2% vs 14.1%, RR, 0.82; 95% CI, 0.77−0.87), and fetal blood sampling rates by 53% (1.75% vs 0.82%, RR, 0.47; 95% CI, 0.38−0.58).[115]

Summary and conclusions

With 6 RCTs, 10 meta-analyses, and more than 20 observational studies published, ST analysis is one of the most intensively studied topics in intrapartum fetal monitoring. Because of differences in study protocols, inclusion criteria, enrollment rates, clinical guidelines, use of fetal blood sampling, and definitions of key outcome parameters as well as inconsistencies in RCT data handling and statistical methodology, the evidence remains difficult to interpret. The vast resources spent on RCTs have failed to guarantee their generalizability to other settings or their ability to reflect everyday clinical practice.

The latest meta-analysis using corrected data from all original RCTs has demonstrated that CTG plus ST significantly reduced fetal blood sampling rates by 41%, operative vaginal delivery rates by 8%, and metabolic acidosis rates by 36%, compared with CTG alone. Similar benefits have not been shown for other intrapartum fetal monitoring techniques. Real-world data from ST analysis use in Europe also show significant reductions in fetal blood sampling and metabolic acidosis rates, without an increase in operative deliveries.

Technical developments and further research must continue. The large amount of data available from previously conducted studies has the potential to provide large data sets of information without the need to conduct new trials. The need to make data accessible for evaluation and reuse has been discussed by leading medical journals[116,117] and by the International Committee of Medical Journal Editors.[118] Effective implementation of this recommendation should prevent the iteration of trials and reduce the costs of technology evaluation. ■

ACKNOWLEDGMENT

Sofia Blad, PhD, Senior Regulatory Consultant Medical Device, Mediteq Svenkebo AB, is greatly acknowledged for assistance related to regulatory aspects. Catherine Arnaud, MD, PhD, Epidemiologist, UMR 1027 INSERM, Team SPHERE, Toulouse III University, Toulouse, France, is also acknowledged for having carefully read the manuscript.

REFERENCES

1. Hon EH, Lee ST. Electronic evaluation of the fetal heart rate. VIII. Patterns preceding fetal death, further observations. Am J Obs Gynecol 1963;87:814–26.
2. Caldeyro-Barcia R. [Mechanisms and significance of the variations registered in the heart rate of the human fetus during labor]. Tokoginecol Pract 1964;23:387–433.
3. Hammacher K. [Electronic control of child birth]. Med Klin 1969;64:1846–51.
4. Ayres-de-Campos D. Electronic fetal monitoring or cardiotocography, 50 years later: what's in a name? Am J Obstet Gynecol 2018;218:545–6.
5. Electronic fetal heart rate monitoring: research guidelines for interpretation. National Institute of Child Health and Human Development Research Planning Workshop. Am J Obstet Gynecol 1997;177:1385–90.
6. National Institute for Health and Clinical Excellence. Intrapartum apartum care for health healthy y women and babies. Nice 2014; p. 1-58. Available at: https://www.nice.org.uk/guidance/cg55. Accessed May 28, 2019.
7. World Health Organization. WHO recommendations: intrapartum care for a positive childbirth experience. Geneva: World Health Organization; 2018.
8. Sabiani L, Le Du R, Loundou A, et al. Intra- and interobserver agreement among obstetric experts in court regarding the review of abnormal fetal heart rate tracings and obstetrical management. Am J Obstet Gynecol 2015;213:856.e1–8.
9. Clark SL, Hamilton EF, Garite TJ, Timmins A, Warrick PA, Smith S. The limits of electronic fetal heart rate monitoring in the prevention of neonatal metabolic acidemia. Am J Obstet Gynecol 2017;216:163.e1–6.
10. Sellier E, Platt MJ, Andersen GL, et al. Decreasing prevalence in cerebral palsy: a multi-site European population-based study, 1980 to 2003. Dev Med Child Neurol 2016;58:85–92.
11. Ayres-De-Campos D, Arulkumaran S. FIGO consensus guidelines on intrapartum fetal monitoring: introduction. Int J Gynecol Obstet 2015;131:3–4.
12. Visser GH, Ayres-de-Campos D. FIGO consensus guidelines on intrapartum fetal monitoring: adjunctive technologies. Int J Gynaecol Obstet 2015;131:25–9.
13. Olofsson P, Noren H, Carlsson A, Rosen KG. Identifying newborns with umbilical cord blood metabolic acidosis by intrapartum cardiotography combined with fetal ECG ST analysis (STAN): comparison of the new and old FIGO systems to classify cardiotocograms. J Matern Fetal Neonatal Med 2018:1–11.
14. Chandraharan E. Fetal scalp blood sampling should be abandoned: FOR: FBS does not fulfil the principle of first do no harm. BJOG An Int J Obstet Gynaecol 2016;123:1770.
15. Clark SL, Paul RH. Intrapartum fetal surveillance: the role of fetal scalp blood sampling. Am J Obstet Gynecol 1985;153:717–20.
16. Westgate J, Greene K. How well is fetal blood sampling used in clinical practice? Br J Obstet Gynaecol 1994;101:250–1.
17. 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. Cochrane Database Syst Rev 2015;2015.
18. Jørgensen JS, Weber T. Fetal scalp blood sampling in labor–a review. Acta Obstet Gynecol Scand 2014;93:548–55.
19. Holzmann M, Wretler S, Cnattingius S, Nordstrom L. Neonatal outcome and delivery mode in labors with repetitive fetal scalp blood sampling. Eur J Obstet Gynecol Reprod Biol 2015;184:97–102.
20. Sabir H, Stannigel H, Schwarz A, Hoehn T. Perinatal hemorrhagic shock after fetal scalp blood sampling. Obstet Gynecol 2010;115(2 Suppl 2, part. 2):419–20.
21. Schaap TP, Moormann KA, Becker JH, et al. Cerebrospinal fluid leakage, an uncommon complication of fetal blood sampling: a case report and review of the literature. Obstet Gynecol Surv 2011;66:42–6.
22. Clark SL, Gimovsky ML, Miller FC. The scalp stimulation test: a clinical alternative to fetal scalp blood sampling. Am J Obstet Gynecol 1984;148:274–7.
23. Holzmann M, Wretler S, Nordstrom L. Absence of accelerations during labor is of little value in interpreting fetal heart rate patterns. Acta Obstet Gynecol Scand 2016;95:1097–103.
24. Skupski DW, Rosenberg CR, Eglinton GS. Intrapartum fetal stimulation tests: a meta-analysis. Obstet Gynecol 2002;99:129–34.
25. Chen H-Y, Chauhan SP, Ananth CV, Vintzileos AM, Abuhamad AZ. Electronic fetal heart rate monitoring and its relationship to neonatal and infant mortality in the United States. Am J Obstet Gynecol 2011;204:491.e1–10.
26. Ananth CV, Chauhan SP, Chen H-Y, D'Alton ME, Vintzileos AM. Electronic fetal monitoring in the United States: temporal trends and adverse perinatal outcomes. Obstet Gynecol 2013;121:927–33.
27. Walsh CA, McMenamin MB, Foley ME, Daly SF, Robson MS, Geary MP. Trends in intrapartum fetal death, 1979—2003. Am J Obstet Gynecol 2008;198:47.e1–7.
28. Shier D, Tilson JL. The Temporal Stage fallacy: a novel statistical fallacy in the medical literature. Med Health Care Philos 2006;9:243–7.
29. Hagberg B, Hagberg G, Beckung E, Uvebrant P. Changing panorama of cerebral palsy in Sweden. VIII. Prevalence and origin in the birth year period 1991—1994. Acta Paediatr 2001;90:271–7.
30. Meberg A, Broch H. Etiology of cerebral palsy. J Perinat Med 2004;32:434–9.
31. Cowan F, Rutherford M, Groenendaal F, et al. Origin and timing of brain lesions in term infants with neonatal encephalopathy. Lancet (London, England) 2003;361:736–42.
32. Hurst JW. Naming of the waves in the ECG, with a brief account of their genesis. Circulation 1998;98:1937–42.
33. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, the American Association for Thoracic Surgery,the Preventive Cardiovascular Nurses Association, the Society for Cardiovascular Angiography and Interventions, and the Society of Thoracic Surgeons. J Am Coll Cardiol 2012;60:e44–164.
34. Scirica BM. Acute coronary syndrome: emerging tools for diagnosis and risk assessment. J Am Coll Cardiol 2010;55:1403–15.
35. Kahn KA, Simonson E. Changes of mean spatial QRS and T vectors and of conventional electrocardiographic items in hard anaerobic work. Circ Res 1957;5:629–33.
36. Southern EM. Fetal anoxia and its possible relation to changes in the prenatal fetal electrocardiogram. Am J Obstet Gynecol 1957;73:233–47.
37. Larks SD, Longo LD. Electrocardiographic studies of the fetal heart during delivery. Obstet Gynecol 1962;19:740–7.
38. Larks SD. The abnormal fetal electrocardiogram: intrauterine fetal difficulty and fetal distress. Obstet Gynecol 1963;22:427–32.
39. Symonds EM. Configuration of the fetal electrocardiogram in relation to fetal acid-base balance and plasma electrolytes. BJOG 1971;78:957–70.
40. Lee KH, Blackwell R. Observations on the configuration of the fetal electrocardiogram before and during labour. BJOG 1974;81:61–9.
41. Davidsen PC. The significance of the foetal electrocardiogram during labour with detailed report of one case. Acta Obstet Gynecol Scand 1971;50:45–9.
42. Pardi G, Tucci E, Uderzo A, Zanini D. Fetal electrocardiogram changes in relation to fetal heart rate patterns during labor. Am J Obstet Gynecol 1974;118:243–50.
43. Hon EH, Lee ST. The fetal electrocardiogram. I. The electrocardiogram of the dying fetus. Am J Obstet Gynecol 1963;87:804–13.
44. Mueller-Heubach E, Adamsons K. Surveillance of the fetus during the intrapartum period. Mt Sinai J Med 1971;38:427–39.
45. Rosén KG, Kjellmer I. Changes in the fetal heart rate and ECG during hypoxia. Acta Physiol Scand 1975;93:59–66.
46. Lindecrantz KG, Lilja H, Widmark C, Rosen KG. Fetal ECG during labour: a suggested standard. J Biomed Eng 1988;10:351–3.

47. Hokegard KH, Eriksson BO, Kjellmer I, Magno R, Rosen KG. Myocardial metabolism in relation to electrocardiographic changes and cardiac function during graded hypoxia in the fetal lamb. Acta Physiol Scand 1981;113:1–7.
48. Hökegård KH, Eriksson BO, Kjellmer I, Magno R, Rosén KG. Myocardial metabolism in relation to electrocardiographic changes and cardiac function during graded hypoxia in the fetal lamb. Acta Physiol Scand 1981;113:1–7.
49. Rosén KG, Dagbjartsson A, Henriksson BÅ, Lagercrantz H, Kjellmer I. The relationship between circulating catecholamines and ST waveform in the fetal lamb electrocardiogram during hypoxia. Am J Obstet Gynecol 1984;149:190–5.
50. Dawes GS, Mott JC, Shelley HJ. The importance of cardiac glycogen for the maintenance of life in foetal lambs and new-born animals during anoxia. J Physiol 1959;146:516–38.
51. Yli BM, Källén K, Stray-Pedersen B. Amer-Wåhlin I. Intrapartum fetal ECG and diabetes. J Matern Neonatal Med 2008;21.
52. Rosen KG, Amer-Wahlin I, Luzietti R, Noren H. Fetal ECG waveform analysis. Best Pr Res Clin Obs Gynaecol 2004;18:485–514.
53. Welin AK, Blad S, Hagberg H, Rosén KG, Kjellmer I, Mallard C. Electrocardiographic changes following umbilical cord occlusion in the midgestation fetal sheep. Acta Obstet Gynecol Scand 2005;84:122–8.
54. Blad S, Welin AK, Kjellmer I, Rosén KG, Mallard C. ECG and heart rate variability changes in preterm and near-term fetal lamb following LPS exposure. Reprod Sci 2008;15:572–83.
55. Lilja H, Greene KR, Karlsson K, Rosen KG. ST waveform changes of the fetal electrocardiogram during labour—a clinical study. Br J Obstet Gynaecol 1985;92:611–7.
56. Lilja H, Karlsson K, Lindecrantz K, Rosen KG. Microprocessor based waveform analysis of the fetal electrocardiogram during labor. Int J Gynaecol Obstet 1989;30:109–16.
57. Arulkumaran S, Lilja H, Lindecrantz K, Ratnam SS, Thavarasah AS, Rosen KG. Fetal ECG waveform analysis should improve fetal surveillance in labour. J Perinat Med 1990;18:13–22.
58. Rosén KG. Device for reducing signal noise in a fetal ECG signal. Available at: https://patents.google.com/patent/US6658284B1. Accessed May 28, 2019.
59. Westgate J, Harris M, Curnow JS, Greene KR. Plymouth randomized trial of cardiotocogram only versus ST waveform plus cardiotocogram for intrapartum monitoring in 2400 cases. Am J Obstet Gynecol 1993;169:1151–60.
60. Amer-Wahlin I, Hellsten C, Noren H, et al. Cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram for intrapartum fetal monitoring: a Swedish randomised controlled trial. Lancet (London, England). 2001;358:534–8.
61. Ojala K, Vaarasmaki M, Makikallio K, Valkama M, Tekay A. A comparison of intrapartum automated fetal electrocardiography and conventional cardiotocography—a randomised controlled study. BJOG 2006;113:419–23.
62. Vayssière C, David E, Meyer N, et al. A French randomized controlled trial of ST-segment analysis in a population with abnormal cardiotocograms during labor. Am J Obstet Gynecol 2007;197.
63. Westerhuis MEMH, Visser GHA, Moons KGM, et al. Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. Obstet Gynecol 2010;115:1173–80.
64. Potti S, Berghella V. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: a meta-analysis of randomized trials. Am J Perinatol 2012;29:657–64.
65. Salmelin A, Wiklund I, Bottinga R, et al. Fetal monitoring with computerized ST analysis during labor: a systematic review and meta-analysis. Acta Obstet Gynecol Scand 2013;92:28–39.
66. Becker JH, Bax L, Amer-Wahlin I, et al. ST analysis of the fetal electrocardiogram in intrapartum fetal monitoring: a meta-analysis. Obstet Gynecol 2012;119:145–54.
67. Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. Cochrane database Syst Rev 2013:CD000116.
68. Schuit E, Amer-Wahlin I, Ojala K, et al. Effectiveness of electronic fetal monitoring with additional ST analysis in vertex singleton pregnancies at >36 weeks of gestation: an individual participant data metaanalysis. Am J Obstet Gynecol 2013;208.
69. Olofsson P, Ayres-De-Campos D, Kessler J, Tendal B, Yli BM, Devoe L. A critical appraisal of the evidence for using cardiotocography plus ECG ST interval analysis for fetal surveillance in labor. Part II: the meta-analyses. Acta Obstet Gynecol Scand 2014;93:571–86.
70. Belfort MA, Saade GR, Thom E, et al. A randomized trial of intrapartum fetal ECG ST-segment analysis. N Engl J Med 2015;373:632–64.
71. Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. Cochrane Database Syst Rev 2015;12:CD000116.
72. 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;127:127–35.
73. Blix E, Brurberg KG, Reierth E, Reinar LM, Oian P. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: a systematic review and meta-analysis of randomized trials. Acta Obstet Gynecol Scand 2016;95:16–27.
74. Malin GL, Morris RK, Khan KS. Strength of association between umbilical cord pH and perinatal and long term outcomes: systematic review and meta-analysis. BMJ 2010;340:c1471.
75. Ayres-De-Campos D, Arulkumaran S. Corrigendum: FIGO consensus guidelines on intrapartum fetal monitoring: physiology of fetal oxygenation and the main goals of intrapartum fetal monitoring (Int J Gynecol Obstet [2015] 131:5-8). Int J Gynecol Obstet 2016;133:255.
76. Hafstrom M, Ehnberg S, Blad S, et al. Developmental outcome at 6.5 years after acidosis in term newborns: a population-based study. Pediatrics 2012;129:e1501–7.
77. Jongsma HW. Introduction to studies within the EC Working Group "Fetal Electro- and Phonocardiography." J Perinat Med 1991;19(Suppl 1):333–6.
78. Rosén KG, Luzietti R. Intrapartum fetal monitoring: Its basis and current developments. Prenat Neonatal Med 2000;5:155–68.
79. Weindling AM. The confidential enquiry into maternal and child health (CEMACH). Arch Dis Child 2003;88:1034–7.
80. Ojala K, Mäkikallio K, Haapsamo M, Ijäs H, Tekay A. Interobserver agreement in the assessment of intrapartum automated fetal electrocardiography in singleton pregnancies. Acta Obstet Gynecol Scand 2008;87:536–40.
81. CORDIS. Dissemination of a knowledge based system for determining appropriate intervention during labour based on a qualified analyses of the foetal electrocardiogram. Available at: https://cordis.europa.eu/project/rcn/52005/factsheet/en. Accessed July 17, 2019.
82. Amer-Wåhlin I, Kjellmer I, Marå ÚI K, Olofsson P, RosÉn KG. Swedish randomized controlled trial of cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram revisited: analysis of data according to standard versus modified intention-to-treat principle. Acta Obstet Gynecol Scand 2011;90.
83. Westerhuis MEMH, Visser GHA, Moons KGM, Zuithoff N, Mol BWJ, Kwee A. Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. Obstet Gynecol 2011;117(2 Pt 1):406–7.
84. Olofsson P, Ayres de Campos D, Kessler J, Tendal B, Yli BM DL. A critical appraisal of the evidence for using cardiotocography plus ECG ST interval analysis for fetal surveillance in labor. Part I: the randomized controlled trials. Acta Obs Gynecol Scand 2014;93:556–68.
85. Gittelsohn J, Shankar A, West K, Ram R, Gnywali T. Estimating reactivity in direct observation studies of health behaviors. Hum Organ 1997;56:182–9.
86. Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 2017;17.

87. Welin A-K, Noren H, Odeback A, Andersson M, Andersson G, Rosen KG. STAN, a clinical audit: the outcome of 2 years of regular use in the city of Varberg, Sweden. Acta Obstet Gynecol Scand 2007;86:827–32.
88. Blix E, Brurberg KG, Reierth E, Reinar LM, Oian P. Statistical significance is not necessarily equal to clinical significance. Acta Obstet Gynecol Scand 2016;95:1192.
89. Bhide A, Chandraharan E, Acharya G. Fetal monitoring in labor: Implications of evidence generated by new systematic review. Acta Obstet Gynecol Scand 2016;95:5–8.
90. Blix E, Brurberg KG, Reierth E, Reinar LM, Oian P. STAN technology, surrogate outcomes and possible sources of bias. Acta Obstet Gynecol Scand 2016;95:608–9.
91. Olofsson P. Belittling of a significant decline in neonatal metabolic acidosis rate achieved by STAN monitoring. Acta Obstet Gynecol Scand 2016;95:604–5.
92. Vayssiere C, Ehlinger V, Paret L, Arnaud C. Is STAN monitoring associated with a significant decrease in metabolic acidosis at birth compared with cardiotocography alone? Review of the three meta-analyses that included the recent US trial. Acta Obstet Gynecol Scand 2016;95:1190–1.
93. Kessler J, Yli BM, Stray-Pedersen B, Jacobsen AF, Saugstad OD, Henriksen T. Why did the authors perform a meta-analysis of studies with primary endpoints they consider clinically unimportant? Acta Obstet Gynecol Scand 2016;95:606–7.
94. Amer-Wåhlin I, Källén K, Herbst A, Rydhstroem H, Sundström AK, Maršál K. Implementation of new medical techniques: Experience from the Swedish randomized controlled trial on fetal ECG during labor. J Matern Neonatal Med 2005;18:93–100.
95. 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;152:524–39.
96. Schuit E. Reply: To PMID 23333546. Am J Obstet Gynecol 2013;209:394–5.
97. Rosen KG, Noren H, Carlsson A. FHR patterns that become significant in connection with ST waveform changes and metabolic acidosis at birth. J Matern Fetal Neonatal Med 2018:1–6.
98. Chen LF, Vander Weg MW, Hofmann DA, Reisinger HS. The Hawthorne Effect in infection prevention and epidemiology. Infect Control Hosp Epidemiol 2015;36:1444–50.
99. Xodo S, Saccone G, Schuit E, Amer-Wåhlin I, Berghella V. Why STAN might not be dead. J Matern Neonatal Med 2017;30.
100. Barnish MS, Turner S. The value of pragmatic and observational studies in health care and public health. Pragmatic Obstet Res 2017;8:49–55.
101. Saturni S, Bellini F, Braido F, et al. Randomized controlled trials and real life studies. Approaches and methodologies: a clinical point of view. Pulm Pharmacol Ther 2014;27:129–38.
102. Kwee A, Van Der Hoorn-Van Den Beld CW, Veerman J, Dekkers AHS, Visser GHA. STAN® S21 fetal heart monitor for fetal surveillance during labor: An observational study in 637 patients. J Matern Neonatal Med 2004;15:400–7.
103. Ross MG, Devoe LD, Rosen KG. ST-segment analysis of the fetal electrocardiogram improves fetal heart rate tracing interpretation and clinical decision making. J Matern Fetal Neonatal Med 2004;15:181–5.
104. Devoe LD, Ross M, Wilde C, et al. United States Multicenter Clinical Usage Study of the STAN 21 electronic fetal monitoring system. Am J Obstet Gynecol 2006;195:729–34.
105. Noren H, Blad S, Carlsson A, et al. STAN in clinical practice—the outcome of 2 years of regular use in the city of Gothenburg. Am J Obstet Gynecol 2006;195:7–15.
106. Kale A, Chong Y-S, Biswas A. Effect of availability of fetal ECG monitoring on operative deliveries. Acta Obstet Gynecol Scand 2008;87:1189–93.
107. Massoud M, Giannesi A, Amabile N, Manevy M, Geron G, Gaucherand P. Fetal electrocardiotocography in labor and neonatal outcome: an observational study in 1889 patients in the French center of Edouard Herriot, Lyon. J Matern Fetal Neonatal Med 2007;20:819–24.
108. Doria V, Papageorghiou AT, Gustafsson A, Ugwumadu A, Farrer K, Arulkumaran S. Review of the first 1502 cases of ECG-ST waveform analysis during labour in a teaching hospital. BJOG 2007;114:1202–7.
109. Colov NSP. [Need for extensive education when implementing new foetal monitoring technology]. Ugeskr Laeger 2007;169:3294–7.
110. Melin M, Bonnevier A, Cardell M, Hogan L, Herbst A. Changes in the ST-interval segment of the fetal electrocardiogram in relation to acid-base status at birth. BJOG 2008;115:1669–75.
111. Rzepka R, Torbe A, Kwiatkowski S, Blogowski W, Czajka R. Clinical outcomes of high-risk labours monitored using fetal electrocardiography. Ann Acad Med Singapore 2010;39:27–32.
112. Noren H, Carlsson A. Reduced prevalence of metabolic acidosis at birth: an analysis of established STAN usage in the total population of deliveries in a Swedish district hospital. Am J Obstet Gynecol 2010;202:546.e1–7.
113. Ragupathy K, Ismail F, Nicoll AE. The use of STAN monitoring in the labour ward. J Obstet Gynaecol 2010;30:465–9.
114. Landman AJEMC, Immink-Duijker ST, Mulder EJH, et al. Significant reduction in umbilical artery metabolic acidosis after implementation of intrapartum ST waveform analysis of the fetal electrocardiogram. Am J Obstet Gynecol 2019;221:63.e1–13.
115. Timonen S, Holmberg K. The importance of the learning process in ST analysis interpretation and its impact in improving clinical and neonatal outcomes. Am J Obstet Gynecol 2018;18:620.e1–7.
116. Krumholz HM, Terry SF, Waldstreicher J. Data acquisition, curation, and use for a continuously learning health system. JAMA 2016;316:1669–70.
117. Najafzadeh M, Schneeweiss S. From trial to target populations—calibrating real-world data. N Engl J Med 2017;376:1203–5.
118. Ross JS. Clinical research data sharing: what an open science world means for researchers involved in evidence synthesis. Syst Rev 2016;5:159.
119. Yin L. What physicians should know about the regulation of obstetric and gynecologic medical devices. J Reprod Med 1983;28:3–11.
120. Food and Drug Administration. Level of assessment. Available at: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?FR=884.2050. Accessed May 28, 2019.
121. Hill D, Stewart D, Balman E. What is ce marking? How technologies are classified, and how to navigate the system. Value Heal J Int Soc Pharmacoeconomics Outcomes Res 2015;18:A367.
122. Food and Drug Administration. Premarket approval. Available at: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P020001. Accessed May 28, 2019.
123. Amer-Wåhlin I, Bördahl P, Eikeland T, et al. ST analysis of the fetal electrocardiogram during labor: Nordic observational multicenter study. J Matern Neonatal Med 2002;12.
124. Food and Drug Administration. Product labeling. Available at: https://www.accessdata.fda.gov/cdrh_docs/pdf2/P020001C.pdf. Accessed May 28, 2019.
125. Food and Drug Administration. Investigational device exemption. Available at: https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/InvestigationalDeviceExemptionIDE/default.htm. Accessed May 28, 2019.
126. Higgins JPT, Lasserson T, Chandler J, Tovey D CR. Methodological expectations of cochrane intervention reviews. London, UK: Cochrane; 2016.
127. Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess the quality of randomized controlled trials: a systematic review. Phys Ther 2008;88:156–75.
128. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement.

Quality of reporting of meta-analyses. Lancet (London, England) 1999;354:1896–900.
129. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62:1013–20.
130. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.
131. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. CMAJ 1997;156:1411–6.
132. Devoe LD, Ross M, Wilde C, et al. United States multicenter clinical usage study of the STAN 21 electronic fetal monitoring system. Am J Obstet Gynecol 2006;195L729—34.
133. Vayssiere C, Haberstich R, Sebahoun V, David E, Roth E, Langer B. Fetal electrocardiogram ST-segment analysis and prediction of neonatal acidosis. Int J Gynecol Obstet 2007;97:110–4.
134. Doret M, Massoud M, Constans A, Gaucherand P. Use of peripartum ST analysis of fetal electrocardiogram without blood sampling: a large prospective cohort study. Eur J Obstet Gynecol Reprod Biol 2011;156:35–40.
135. Kessler J, Moster D, Albrechtsen S. Intrapartum monitoring of high-risk deliveries with ST analysis of the fetal electrocardiogram: an observational study of 6010 deliveries. Acta Obstet Gynecol Scand 2013;92:75–84.

Supplementary material

Regulatory approval processes

Medical device development is heavily regulated in most of the world. In the United States, there are 3 categories of devices, depending on the device's intended use and the risks associated with its use[119]: a 510(k) premarket notification in which the manufacturer must demonstrate equivalence of the new device to one that is legally accepted and in commercial distribution in the United States before May 28, 1976 (class I); or a device that has been determined by Food and Drug Administration (FDA) to be substantially equivalent (class II); or devices that either pose a significant risk of illness or injury (high-risk devices) or are not substantially equivalent to class I or II devices and require premarket approval (PMA) (class III).

The PMA process is more complex and includes the submission of clinical data to support claims made for the device. The higher the classification, the greater the level of assessment required.[120] STAN has received PMA as a class III device from the FDA.

In Europe, the manufacturer's declaration that the product meets the requirements of the applicable EC directives is called Conformité Européenne (CE) certification. This is a mark that indicates conformity with health, safety, and environmental protection standards for products sold within the European Union. The classification that STAN received for its CE mark is the second highest available in Europe (IIB)[121] and indicates that the product was assessed before being placed on the market and meets the requirements of the applicable EC directives.

During the multiyear United States premarket approval process, US users demonstrated their consistency in managing the CTG plus ST data clinically.[122] The clinical guidelines included in the concept were based on guidelines (Table 2), tested in a Nordic study.[123] They informed health care professionals when to intervene in case of ST events, in combination with CTG features (Figure 7).

Once the standards, including guidelines, were approved, they became part of the product labeling and a legally binding prerequisite for the use of ST analysis.[124] STAN was approved in Europe and in the United States after long premarket approval processes. According to regulations, any change to the clinical use, including a different use in a trial, must be approved by the FDA though an Investigational Device Exemption (IDE) process, in a randomized controlled trial under the responsibility of the sponsor.[125]

Meta-analysis methodology

First, the studies to be aggregated for each criterion (outcome) must be chosen. Because this critical choice is based on a subjective assessment, it is recommended that it be conducted independently by 2 reviewers.[126] The final decision depends on evaluations covering 2 areas: (1) how to identify the studies on the research topic (literature search strategy, list of key terms, study eligibility criteria such as year of publication, language, etc), and (2) which studies should be included in the meta-analysis, which requires assessment of the quality and homogeneity of the study designs (eligibility criteria for participants, interventions being compared, definition and method for assessing the outcome, risk of bias, blinding, etc).

Numerous checklists and scales have been proposed to help these assessments. For example, Olivo et al[127] found 21 scales assessing the quality of randomized controlled trials in medical research.[127]

The second step involves the data extraction. The reviewers must check whether all the data needed to compute an aggregated estimation of the intervention effect are available. If necessary, they may contact authors to obtain supplemental data or information and must verify whether revised data were published.

The statistician should then consider the following issues: choosing a suitable effect measure (for example, risk ratio vs odds ratio or risk difference), estimating and testing for heterogeneity between studies (3 common parameters can be used, each potentially leading to different decisions and without clear and recognized cutoffs), weighting the studies' results (several weighting scheme exist), and evaluating the presence of publication bias with a funnel plot.

These choices can be different in each meta-analysis process and can lead to varying or even divergent results. Subjectivity in meta-analyses is well recognized, and various checklists have been proposed to assess the methodology. For example, the QUOROM (QUality Of Reporting Of Meta-analysis) statement[128] as well as the AMSTAR (Assessing the Methodological quality of Systematic Reviews) checklist[129] and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement that have replaced the QUOROM statement[130] today can be used for critical appraisal of published systematic reviews. A guide and an algorithm to interpreting discordant systematic reviews have even been published.[131]

Principal stages in meta-analysis: steps in the process involving subjectivity

Study selection
Reproducible systematic search of studies meeting eligibility criteria
Qualitative appraisal of the studies (studies to be omitted from the review).
Inclusion/exclusion/criteria of studies for the statistical analysis (evaluation of homogeneity between the study designs).

Data extraction
Check the definition, measurement, and population considered for the outcome.
Check for revised data.
Human error.

Statistical analysis
Quantification of heterogeneity.
Recognition of heterogeneity (random- or fixed-effect analysis).
Choice of a suitable effect measure (for example, risk ratio).
Handling of missing data.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Appendix A

Summary chart of prior clinical studies

List and comments on clinical trials leading up to the Swedish RCT[1]

Reference: Arulkumaran S, Lilja H, Lindecrantz K, et al.[57]

Observational study: IRB approval, informed consent

Objectives/protocol	Methods/materials	Patient population	Results	Comments
Objective The objective of the study was to evaluate ST waveform changes during labor, with, as the reference, a recording technique using maternal skin. ST waveform changes in the form of T/QRS ratios were correlated with FHR changes, Apgar scores, and cord artery acid-base data. Clinical management FHR formed the basis for management of labor. The STAN recorder was used together with a standard CTG recorder. CTG classification FHR recordings were classified as normal, suspicious, or abnormal, in accordance with FIGO News 1987 recommendations. ST waveform assessment Baseline and increases for more than 5 minutes in T/QRS ratio were considered. The upper level of normal was set at a T/QRS ratio of 0.25. Negative T-wave components were considered abnormal. Respiratory acidemia Cord artery pH <7.15, standard bicarbonate >15 mmol/L. Metabolic acidosis pH <7.15 and standard bicarbonate <15 mmol/L. Perinatal asphyxia Cord artery pH <7.15 and standard bicarbonate <15 mmol/L. Apgar score <4 at 1 minute and <7 at 5 minutes.	ST analyzer STAN 8801 Fetal ECG was obtained from a single spiral electrode with the reference point placed on the maternal thigh. The signal was fed via a preamplifier to a purpose-built microprocessor for automatic assessment of the ST waveform. Average T and QRS amplitudes were calculated from 10 ECGs. The resultant quotient was calculated and plotted in analog form. The ECG signal was sent simultaneously to a FHR monitor and STAN.	Patients in labor with gestational age >34 weeks. Patients with antenatal risk factors were preferentially included. 201 fetuses in vertex presentation were included.	Signal quality Signal quality enabled T/QRS measurements at least every 2 minutes in 77% of the recordings. In the remaining 23%, there were blank periods in the recordings, but because the T/QRS ratio was found be stable, especially during the first stage of labor, these recordings were also included in the analysis. Operative interventions 27 (13%) operative deliveries for fetal distress. 14 with abnormal FHR leading to 8 CD and 6 forceps deliveries. 13 with suspicious FHR leading to 6 cesarean and 7 forceps deliveries. 11 of the 27 had T/QRS changes with average ratios >0.25. 3 of the 27 had cord artery pH <7.15; all 3 were identified by elevated T/QRS ratios. Neonatal outcome 3 babies had clinical evidence of asphyxia: 1 had abnormal FHR tracings and was delivered by cesarean, and the other 2 had suspicious FHR tracings. All 3 cases had T/QRS elevations >0.25. 5 babies had metabolic acidosis. All 5 had T/QRS ratios >0.25 during the first stage of labor. 8 babies had respiratory acidemia: 5 had abnormal T/QRS ratios at the beginning of the last stage, 1 had no ST changes (pH, 7.14), 2 had increased T/QRS ratios 20 minutes before delivery. In 1 case, a 7 minute bradycardia episode occurred (FHR <100 beats).	The available data showed that: It is possible to use the described STAN system for on-line recording of the unfiltered fetal ECG waveform and to analyze the ST waveform. Acute hypoxia was recognized by a rapid rise in the T/QRS ratio. Changes in the ST interval are related to metabolic events in the myocardium and FHR changes to neurophysiological responses during hypoxia. That is, 2 different parameters are involved, involving different responses to fetal hypoxia. The data demonstrated the value of combined studies of FHR and ST waveform analysis. T/QRS ratio is less sensitive for identifying respiratory acidosis, which is not, however, a major threat to the baby. T/QRS ratio <0.25 identifies with 99.3% accuracy a fetus with normal buffering capacity, independent of FHR tracing. Unnecessary cesarean deliveries could thereby be avoided. Because of the low rate of perinatal asphyxia and early interference based on FHR tracings, further studies are required to evaluate whether fetal ECG waveform analysis will improve the diagnosis of perinatal asphyxia.

CD, cesarean delivery; *CTG*, cardiotocography; *ECG*, electrogram; *FHR*, fetal heart rate; *FIGO*, International Federation of Gynecology and Obstetrics; *IRB*, institutional review board; *STAN*, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Reference: Maclachlan NA, Spencer JAD, Harding K, Arulkumaran S. Fetal acidemia, the cardiotocograph and the T/QRS ratio of the fetal ECG during labor. Br J Obstet Gynaecol 1992;99:25-31.

Observational study: IRB approval, informed consent

Objectives/protocol	Methods/materials	Patient population	Results	Comments
Objective The objective was to compare the sensitivity and the positive predictive value of CTG and elevated T/QRS ratio in fetuses with acidemia during labor and at delivery. FBS pH <7.20 and cord artery pH <7.12 were correlated with T/QRS ratio >0.28 and CTG classification. Clinical management T/QRS data were not used for clinical management. FBS was collected when clinically indicated, usually to assess FHR changes. ST waveform assessment The mean of 3 T/QRS ratios before FBS (within 10 minutes) or delivery (within 30 minutes) was taken as representative for comparison with pH or neonatal outcome. CTG classification FHR recordings were classified as normal, suspicious, or abnormal in accordance with FIGO News 1987 recommendations. Acidemia Cord artery pH <7.12 Metabolic acidosis Cord artery pH <7.12, BD_{ecf} >12.0 mmol/L.	ST analyzer STAN 8801 FHR was monitored with a Copeland scalp electrode. The same electrode was used for the ECG data to a STAN monitor for online calculation and recording of the T/QRS ratio. Maternal skin was used for reference. Average T and QRS amplitudes were calculated from 30 ECGs. The resultant quotient was calculated and plotted in analog form. CTG recorders HP 8031 FM6	113 women in labor with term fetuses (37−42 weeks). Pregnancies with abnormal CTG were preferentially included.	Signal quality Interpretable CTG traces within 30 min of labor were available for 93 cases. The quality of ST recordings deteriorated during the second stage of labor. CTG vs other factors The group with suspicious and abnormal CTG showed no increase in the rate of operative deliveries or T/QRS ratio >0.28 but did show a significantly lower median cord artery pH ($P <$ .03 and $P <$ .001, respectively). Neonatal outcome 17 fetuses were classified as acidemic and the median T/QRS ratio in this group was significantly higher than in the nonacidemic group (0.21 and 0.13, respectively, $P <$ .001). 13 of the 17 acidemic babies had metabolic acidosis. CTG results 13 of the acidemic babies had abnormal CTG tracings, for a positive predictive value of 38% and sensitivity of 76%. ST analysis results The positive predictive value for acidemia diagnosed solely from elevated T/QRS ratio was 71% and the sensitivity was 29%. T/QRS ratio did not correlate with scalp pH, but in combination with an abnormal CTG, scalp pH values <7.25 were strongly correlated with rising T/QRS ratios (r = 0.71, $P >$ .001).	These data confirm that CTG interpretation is inaccurate in clinical practice. Interpretation of CTG changes might be improved if the tracings without an associated T/QRS rise could be reliably discounted as false positives. Much larger clinical experiments than currently available are required before T/QRS ratio can be considered for introduction into clinical practice. The T/QRS ratio should be compared with the neonatal outcome. If they are associated, this ratio could be used to reduce the number of false positives and thereby the intervention rate.

Comments related to the paper: the study by Maclachlan et al about 113 term pregnancies used the Copeland scalp clip, which has been shown to adversely affect not only the quality of the ECG signal but also the presentation of the different frequency components contained within the ECG. Furthermore, the study did not discuss the lag time between the end of the STAN recording and delivery, nor did it assess ST waveform configurations other than elevated T waves. The study showed the limitations of using only T/QRS ratio and focusing on first events and trying to relate these to the outcome of labor. Moreover, it did not contain any true case of intrapartum asphyxia (ie, cord artery pH <7.05 and extracellular fluid base deficit >12 mmol/L). It is an illustration of the need for a strict protocol for studies to evaluate the clinical potential of ST waveform analysis. *BD_{ecf}*, BD in the extracellular fluid; *CTG*, cardiotocography; *ECG*, electrogram; *FBS*, fetal blood sampling; *FHR*, fetal heart rate; *IRB*, institutional review board; *STAN*, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Reference: Murphy KW, Russell U, Johnson P, Valente J. Clinical assessment of fetal electrocardiogram monitoring in labor. Br J Obstet Gynaecol 1992;99:32-37.

Observational study: IRB approval, informed consent

Objectives/protocol	Methods/materials	Patient population	Results	Comments
Objective The objective was to investigate the potential of ST waveform analysis during labor. Clinical management Observational study. ST waveform assessment Mean T/QRS ratio for up to 1 hour, at each cervical dilation point of 4, 8, and 10 cm, was compared with CTG changes and with indicators of infant condition at birth: Apgar scores, cord artery acid base, and need for neonatal care. T/QRS ratio >0.25 was considered outside the normal range. CTG classification The CTG was classified as normal according to the Steer criteria (1989). Respiratory acidemia Cord artery pH <7.12. Metabolic acidosis Cord artery pH <7.12, BD_{ecf} >12 mmol/L, Apgar scores <7 at 1 minute.	ST analyzer STAN 8801 prototype FHR was monitored by using a standard scalp electrode. The same electrode was used for the ECG data to a STAN monitor for on-line calculation and recording of the T/QRS ratio. Maternal skin was used as reference. Average T and QRS amplitudes were calculated from 30 ECGs. The resultant quotient was calculated and plotted in analog form.	86 high-risk pregnancies, with a high frequency of inductions, instrumental deliveries, etc.	Signal quality Of the 86 recordings, 3 were excluded because of poor signal quality. Neonatal outcome 7 infants had respiratory acidemia; 4 had clinically significant metabolic acidosis. The most severely affected fetus was identified by T/QRS ratio >0.25 early in labor. During the last stage of labor, the T/QRS ratio decreased, and biphasic ST segments were identified. One of the fetuses showed a T/QRS increase but not >0.25. One was disconnected from ST analysis before the CTG tracing started to deteriorate. The last fetus, with only weak signs of metabolic acidosis, was not identified by CTG or ST tracing. Correlations No significant relation was found between the 1 hour mean T/QRS and Apgar <7 at 1 minute or cord artery pH <7.12. There was a statistical significant correlation between 1 hour mean T/QRS and BD_{ecf} ($r = 0.31$, $n = 39$, $P < .05$).	The mean 1 hour T/QRS ratio is not a good predictor of low Apgar scores, but it did identify the most severely asphyxiated child. This case also indicates that it might be necessary to identify other changes in the ST interval besides the T/QRS ratio. If fetal ECG analysis has a role in future fetal intrapartum surveillance, it might be in conjunction with CTG. Then ST analysis could be used to decrease unnecessary interventions.

BD_{ecf}, BD in the extracellular fluid; *CTG*, cardiotocography; *ECG*, electrogram; *FHR*, fetal heart rate; *IRB*, institutional review board; *STAN*, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Reference: Westgate J, Harris M, Curnow JS, Greene KR.[59]

Interventional study: IRB approval, informed consent

Objectives/protocol	Methods/materials	Patient population	Results	Comments
Objective The objective was to investigate whether CTG plus ST analysis could improve the predictive value of intrapartum surveillance compared with CTG only and could thereby decrease interventions without increased risk for the babies. This prospective clinical trial was divided in a CTG-only and a CTG plus ST arm. The tracings were compared with neonatal outcome and cord acid-base data, both at delivery and retrospectively. CTG arm Interpretation of CTG and management followed accepted clinical guidelines, including FBS option. In the second stage, acutely emerging ST waveform changes for >5 minutes were considered significant. CTG plus ST arm CTG was classified according to the same clinical guidelines as in the CTG arm, but clinical management was modified based on T/QRS ratio and ST segment waveform changes. Metabolic acidosis Cord artery pH <7.05 and BD_{ecf} >12 mmol/L. Birth asphyxia Cord artery pH <7.05, BD_{ecf} >12 mmol/L, Apgar scores <7 at 5 minutes. Active resuscitation for ≥4 minutes and problems in the postnatal period.	ST analyzer STAN 8801 Fetal ECG was recorded with a standard single spiral scalp electrode. Maternal skin was used as reference. The ECG signal was used for on-line calculation and recording of the T/QRS ratio based on averaged 30 beat ECG waveforms. Biphasic/negative ST waveforms were assessed by visual analysis of printed ECG averages. CTG recorder HP 8040A Personnel training All personnel were trained both in CTG classification and ST analysis before and during the study.	2400 high-risk pregnancies after >34 weeks' gestation with no gross fetal abnormality were included. Allocation to intervention was decided by draw of a sealed envelope. CTG arm 1212 cases CTG plus ST arm 1188 cases	Signal quality In 12 cases, ST waveform analysis could not be assessed because of poor signal quality. Intervention rates There was a 46% reduction ($P < .001$) in operative deliveries for fetal distress without an increase in operative deliveries for other reasons. The different CTG patterns were equally distributed in both arms. Neonatal outcome There were more cases of metabolic acidosis (13 vs 5) and more low Apgar scores (<7) (32 vs 20) in the CTG arm. Although this was not statistically significant, it did show a trend toward improved short-term neonatal outcome in the ST plus CTG arm. There were (4 vs 3) cases of birth asphyxia. Negative ST waveforms There were 6 cases identified with persistent negative ST waveforms in association with an abnormal CTG trace. All were depressed at birth and required resuscitation. These 6 included 2 of the 3 cases of birth asphyxia in the CTG plus ST arm.	The study showed that ST waveform analysis discriminates CTG changes in labor and that the protocol was safe. The results from this study confirm that the ST plus CTG analysis significantly reduces interventions without adverse effects on neonatal outcome. Further studies are required for statistical confirmation of the trend toward fewer cases of metabolic acidosis in the CTG plus ST arm. The retrospective analysis of data supports the trend in improved neonatal outcome. CTG plus ST analysis does not require additional procedures, provides continuously available information, and is therefore more likely to aid in appropriate and timely decision making than FBS.

Comments related to the paper: a retrospective analysis of the CTG showed operative deliveries for fetal distress in 2.7% of cases with normal CTG in the CTG-only group, compared with 0.3% in the STAN group. Intermediate CTG patterns led to operative interventions in 19.5% of CTG-only and 9.6% of CTG plus ST cases, and an abnormal CTG the intervention rate to interventions for 44.4% and 35.3%, respectively. In the CTG-only arm, 43% of operative interventions were judged unnecessary in the CTG arm, compared with 5% in the STAN arm of the trial. There were no significant differences in the measures of neonatal outcome, but fewer low 5 minute Apgar scores and fewer cases of metabolic acidosis in the ST plus CTG arm were apparent; there was also a significant reduction in the use of FBS.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Interventions should have been performed for 18% of abnormal traces in the CTG arm (2 cases of asphyxia) and 9% in the STAN arm (1 case of asphyxia). Three patterns of ST plus CTG change occurred:

1. ST persistently raised, CTG normal
   In this group, the mean cord artery pH of 7.28 was significantly higher than in any other group. We believe the slightly raised ST waveform reflects sympathoadrenal stimulation from the general arousal of labor; the neonatal outcome in these cases was excellent.
2. ST waveform rising, CTG abnormal and deteriorating
   The ST waveform rose and the CTG deteriorated and became abnormal in a group of fetuses with significantly lower mean cord artery pH (7.05; 7.02−7.08) and higher base deficit (7.6 mmol/L; 6.1−9.1) than in any other groups but still had normal outcomes. We believe this pattern represents fetuses who were developing a metabolic acidosis as a result of significant hypoxia.
3. ST Segment depressed with biphasic/negative T waves
   The ST waveform was negative or the ST segment depressed in a small number, but all of these cases were depressed at birth and required resuscitation and had low arterial pHs (<7.08) and BD_{ecf} >10 mmol/L (when available). Similar cases have been reported from other groups with metabolic acidosis, growth retardation, or asphyxiation death. These patterns are therefore entirely consistent with the animal data and do suggest a possibility of distinguishing the normal fetus suffering acute hypoxia, showing ST elevation and high T waves, from the chronically hypoxemic fetus who undergoes further acute hypoxic insult and shows negative/biphasic ST waveforms. The 3 fetuses in the CTG plus ST arm who were clinically affected all had ST events that were not recognized by the operator. This finding focused development on automatic assessment of ST events and the new STAN S21 unit was thus designed on the basis of the experience during the Plymouth RCT. BD_{ecf}, BD in the extracellular fluid; *CTG*, cardiotocography; *ECG*, electrogram; *FBS*, fetal blood sampling; *IRB*, institutional review board; *STAN*, ST segment analysis.

Benefits from ST waveform monitoring:

1. The ST waveform provides another physiological variable from the same scalp electrode used to obtain the fetal heart rate.
2. ST waveform change reflects the metabolic events occurring at a tissue level in response to compensatory mechanisms for oxygen lack in a vital central organ. All the evidence from animal data and human studies so far suggest these changes occur before there is any tissue damage.
3. The use of the CTG alone results in many unnecessary and inappropriate interventions. The sensible use of ST waveform in combination with CTG results in a significant and safe reduction in such interventions.
4. The physiology of ST waveform change is better understood than fetal heart rate changes, and its use is a good way to introduce clinicians to the complex physiological responses that occur in labor and thereby improve their interpretation of events as they affect the fetus.

Risks of ST waveform monitoring:

1. High-quality signals are needed for ECG analysis and good application of a single spiral fetal scalp electrode is required. Signal noise may give erroneous T/QRS ratio results.
2. Adequate education of staff in the concepts of both ST waveform and CTG analysis is essential for correct clinical interpretation.
3. There is currently too much emphasis on the T/QRS ratio, which has the benefit of being quantifiable but is only one aspect of ST waveform assessment. It is no surprise that studies attempting to correlate T/QRS values with cord artery pH across the normal range of both parameters find almost no relationship at all. Important changes in the ST segment such as ST depression may be missed if the whole waveform is not examined. This examination also assesses signal quality and checks that T/QRS measurements are not erroneous as a result of noise.

Reference: Luzietti R, Erkkola R, Hasbargen U, Mattsson LA, Thoulon JM, Rosen KG. European Community multi-center trial, "Fetal ECG Analysis During Labor": ST plus CTG analysis. J Perinat Med 1999;27:431-40.

Observational study: IRB approval, informed consent

Objectives/protocol	Methods/materials	Patient population	Results	Comments
Objective The objective was to identify changes in the fetal ECG waveform in cases of verified fetal hypoxia. In this study, the main focus was on changes in the T/QRS ratio, assessed by an automatic system for trend analysis of ST changes together with automatic identification of ECG complexes with ST segment abnormalities (ie, ST segment depression). Clinical management In this prospective study, the ECG waveform information was not available to the clinician during delivery. Retrospective assessment The CTG plus ST tracings were assessed retrospectively, with assessors blinded to the clinical outcome. The data were grouped according to the categories of action of the CTG plus ST clinical guidelines. Outcome parameters The outcome parameters considered were: birthweight; Apgar scores at 1, 5, and 10 minutes; cord artery and vein acid-base assessment; need and method of resuscitation; and transfer to neonatal intensive care unit. ST waveform assessment The ST changes considered were: episodic T/QRS rise (>0.10 for <10 minutes), T/QRS baseline rise (>0.05 for >10 minutes) with repeated biphasic STs, and appearance of repeated negative T waves with ST depression. ST waveform changes were assessed differently, depending on the CTG classification.	ST analyzer STAN 8801 recorder connected to a PC for further signal processing, with data reduction and storage. The data were further processed to regenerate a CTG trace and a 30 beat ECG average for ST waveform analysis. The off-line signal checked for signal quality to ensure that only high-quality ECG waveforms were included. T/QRS ratio was automatically calculated and ST segments with negative slopes (biphasic ST) were identified. ST waveform changes were identified both through visual inspections of the CTG plus ST traces and through an automatic PC-based algorithm, the ST log.	618 cases were recorded, but because of data collection inconsistencies, only 320 cases could be reviewed and compared with the original case notes. All cases included had a gestational age >36 weeks.	Signal quality. The quality of the traces allowed 84% of the available ECG to be used for FHR analysis and 80% of those for ST analysis. Neonatal outcome This data set included 6 cases of intrapartum hypoxia. All were identified by ST events. One additional case had cerebral bleeding, probably associated with ventouse delivery for failure to progress. CTG analysis On retrospective analysis, the CTG was abnormal in 55 cases. ST waveform analysis Baseline rise in the T/QRS ratio occurred in 5 cases, all associated with abnormal CTG. All of these neonates had evidence of intrapartum hypoxia. Episodic T/QRS rise occurred in 16 cases, all but 1 associated with an abnormal CTG. All babies had an uneventful neonatal period. In 5 cases, biphasic STs were intermittent, brief, and associated with a normal CTG. All these cases had a normal outcome. The 1 baby with persistent biphasic STs and an abnormal CTG had evidence of intrapartum hypoxia. Operative deliveries 30 cases of operative delivery: 18 instrumental vaginal and 12 emergency cesareans. Operative delivery for fetal distress was performed for only 2 of the 6 hypoxic babies.	The strong association between ST waveform changes and adverse intrapartum events is illustrated by the fact that 6 of 6 cases with evidence of intrapartum asphyxia showed ST changes. At the same time, no action was taken in 4 of the cases with the most marked asphyxiation. Thus, there is little doubt that ST waveform analysis may add to current techniques for intrapartum fetal surveillance. The clinical guidelines used are based on the combined CTG plus ST analysis. The latter parameter allows for a more detailed assessment of adverse events in labor associated with hypoxia. A new STAN recorder containing the ST log function is to be tested in a second randomized controlled trial. This study should have power enough to show the degree to which perinatal outcome can be improved by using CTG plus ST analysis.

CTG, cardiotocography; *ECG*, electrogram; *IRB*, institutional review board; *STAN*, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Reference: Amer-Wåhlin I, Bördahl P, Eikeland T, et al.[123]

Observational study: IRB approval, informed consent

Objectives/protocol	Methods/materials	Patient population	Results	Discussion
Objective Because considerable improvements in signal processing took place after data collection for the EC multicenter trial was finalized, another retrospective observational study was conducted to identify changes in the ST waveform. The accuracy of the STAN clinical guidelines for intervention and the new ST log function was also tested. In 3 of the centers participating, the trial became part of the preparation for the Swedish RCT. Clinical management Observational study, the ST data were available to the clinician, but clinical action was based on standard procedures. Retrospective analysis The CTG plus ST traces were assessed retrospectively and blinded to the clinical outcome. The data were grouped according to the categories of action recommended by the CTG plus ST clinical guidelines. Assessment of the clinical outcome was based on cord artery and vein acid-base data, Apgar scores, need for resuscitation, referral to neonatal intensive care, and signs of abnormal neonatal neuromuscular findings. ST waveform assessment The ST changes considered were episodic T/QRS rise ($>$0.10 for $<$10 minutes), T/QRS baseline rise ($>$0.05 for $>$10 minutes), repeated biphasic STs and appearance of repeated negative T waves with ST depression. ST waveform changes were assessed differently, depending on the CTG classification.	ST analyzer The prototype of the STAN S 21 (STAN ESST) ST waveform changes were identified both through visual inspections of the CTG plus ST tracings and through an automatic PC-based algorithm called ST log.	574 deliveries with gestational age $>$36 weeks.	Neonatal outcome 15 cases were identified as exposed to intrapartum hypoxia: 5 of them had neonatal neurological symptoms. All 5 were identified as abnormal cases according to CTG plus ST clinical guidelines during the first first stage of labor. The other 10 babies had metabolic acidosis only (cord artery pH $<$7.05 and BD_{ecf} $>$12 mmol/L). 2 had changes in the first stage of labor and the remaining 8 showed ST changes during the second stage of labor. 12 of the ST events were T/QRS baseline increases. One case had an episodic T/QRS increase, another consistent ST depression with negative T waves, and the final case a preterminal CTG as a predominant finding with 1 episodic T/QRS rise. 8 cases had cord artery acidemia only (pH $<$7.05 but BD_{ecf} $<$12 mmol/L). They were all unaffected at birth. 7 of them had shown CTG plus ST abnormalities. Clinical guidelines The sensitivity of CTG plus ST clinical guidelines for recommending intervention was 100% (15/15) for cases with neonatal symptoms and/or metabolic acidosis and 95.8% (22/23) when the respiratory acidosis cases were included. The corresponding figures for the specificity were 95.0% and 96.4%. Operative interventions Had the CTG plus ST clinical guidelines been followed, the operative intervention rate would have been 7.5%, compared with the actual rate of 15.3%.	The Plymouth trial showed that cases with ST elevation and abnormal CTG all had cord artery pH $\leq$7.15. In this study, 86% of the cases in which the STAN clinical guidelines called for intervention had cord artery pH $\leq$7.15. The difference may be explained by the improvements in signal quality and the ability of the ST log to identify ST changes and biphasic ST patterns more accurately at an earlier stage of hypoxia. The experience gained thus far demonstrates the ability of CTG plus ST clinical guidelines, supported by computerized assessment of ST changes to identify babies at risk of intrapartum hypoxia. Further progress of the STAN concept will depend on the outcome of the Swedish multicenter randomized controlled trial.

BD_{ecf}, BD in the extracellular fluid; *CTG*, cardiotocography; *EC*, Conformité Européenne; *IRB*, institutional review board; *STAN*, ST segment analysis.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

Appendix B

Cases with adverse outcome in Swedish RCT[60]

Interval between onset of FHR and ST changes and time of delivery, related to perinatal outcome and metabolic acidosis in cases with adverse/complicated neonatal outcome

			Interval between onset of changes and delivery (min)				
Case N:	Outcome	Metabolic acidosis	Intermediary CTG	Abnormal CTG	Preterminal CTG	ST events	Comments
MAB 215	Perinatal death	Yes		184			Disconnected for 97 minutes
LDG 245**	Perinatal death	Yes			66		Preterminal CTG from onset
MAE 473	Perinatal death	Yes	114	62	46	59	Episodic ST rise of 0.11, in conjunction with baseline FHR 170 bpm followed by a prolonged deceleration of 6 minutes' duration
OEK 309	Moderate/severe encephalopathy	Yes	740	680		740	Baseline T/QRS rise 0.11 plus episodic rise of 0.18 in conjunction with repeated prolonged variable decel (>60 seconds)
LDA 258	Moderate/severe encephalopathy	Yes	142	21		103	Episodic T/QRS of 0.11, 103 minutes in conjunction with baseline FHR 160 plus prolonged and marked variable deceleration (>60 seconds; >60 bpm)
MAA 479	Moderate/severe encephalopathy	Yes	600	477		428	Continuous biphasic ST, 428 minutes in association with baseline FHR 190 bpm followed by baseline T/QRS of 0.07 after a deceleration of 5 minutes' duration of 230 minutes before delivery. Maternal pyrexia.
MAC 516	Moderate/severe encephalopathy	Yes		174	14	149	Episodic T/QRS 0.16, 149 min in conjunction with a prolonged decel of 3 min duration + Baseline FHR of 180 bpm followed by baseline T/QRS of 0.08 last 39 minutes, associated with baseline FHR; 160 bpm and prolonged variable decel (> 60 seconds' duration)
OEH 379	Moderate/severe encephalopathy	No		494			Shoulder dystocia, normal ST
LDA 23	Moderate/severe encephalopathy	No	234	184	184	141	6 episodes of biphasic ST associated with baseline FHR 160–170 bpm and no reactivity/variability
MAD 408	Moderate/severe encephalopathy	No			47		Preterminal FHR from onset, baseline FHR 150 bpm; no reactivity/variability, normal ST
OEK 394	SCBU+	Yes	72	32		22	Prolonged variable decel plus baseline FHR of 160 bpm during second stage. Baseline T/QRS rise of 0.07 last 22 minutes
OEJ 330	SCBU	Yes	301	171	11	23	Baseline T/QRS rise of 0.07 last 23 minutes in association with baseline FHR of 160 bpm and prolonged variable deceleration.
LDC 364	SCBU	Yes	167	127	107	88	Episode of biphasic ST, 88 minutes before delivery in association with prolonged deceleration/lack of variability and reactivity, baseline FHR of 150–160 bpm
LDA 275	SCBU	Yes	45	23		15	Baseline T/QRS of 0.07 last 15 minutes in association with a prolonged deceleration (4 minutes) with onset active pushing followed by repeated prolonged deceleration lasting 1–3 minutes, baseline FHR 120 bpm.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019. *(continued)*

Interval between onset of FHR and ST changes and time of delivery, related to perinatal outcome and metabolic acidosis in cases with adverse/complicated neonatal outcome *(continued)*

			Interval between onset of changes and delivery (min)				
Case N:	Outcome	Metabolic acidosis	Intermediary CTG	Abnormal CTG	Preterminal CTG	ST events	Comments
LDA 298	SCBU	Yes	98	65		27	Baseline T/QRS of 0.06 last 23 minutes in association with bradycardia episodes of 2–3 minutes' duration in second stage. Baseline FHR 120 bpm.
OEF 305	SCBU	Yes	467	167		461	Baseline T/QRS of 0.09, 2 hour episode starting 461 minutes before delivery in association with baseline FHR 150 bpm and loss of reactivity/ variability for 80 minutes.
OEB 363	SCBU	Yes	54			54	Baseline T/QRS of 0.10 last 74 minutes in association with baseline FHR 160 bpm and prolonged variable deceleration.
OEH 371	SCBU	Yes		216		212	Episodic T/QRS (0.11) 212 minutes before delivery in association with prolonged and marked variable deceleration followed by baseline T/QRS last 70 minutes, associated with baseline FHR 150–160 bpm and marked variable deceleration (>60 beats lost).
MAD 342	SCBU	Yes	—	214			Poor signal quality in second stage, no ST events.
MAD 438	SCBU	Yes	—	—			No ST or FHR information available during the last 60 minutes.
OEB 282b	SCBU	Yes	338	261		231	6 biphasic ST in association with baseline FHR of 180 bpm and prolonged deceleration followed by baseline T/QRS of 0.08 at 86 minutes, associated with marked variable deceleration, baseline FHR; 150 bpm.
OEH 330	SCBU	Yes	185	119		119	Baseline T/QRS last 119 minutes is associated with repeated bradycardia episodes of up to 7 minutes' duration in second stage.
OEB 239	SCBU	Yes	49	39		33	Baseline T/QRS rise 0.06–0.21 last 33 minutes in association with saltatory FHR followed by persistent bradycardia (<100 bpm for 10 minutes) with onset of active pushing
MAE 251	SCBU	Yes	665	655			Normal ST at start, no ST information available last 236 minutes, monitored by conventional CTG monitor
MAD 494	SCBU	Yes		17	7	10	Baseline T/QRS rise of 0.25 during last 10 minutes in association with persistent bradycardia.
LDD 263	SCBU	No	182				Delivery trauma, no ST events, no ST data last 24 minutes
MAA 293	SCBU	No		45		32	Baseline T/QRS rise of 0.06 in association with prolonged bradycardia (3 minutes and longer)

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019. *(continued)*

Interval between onset of FHR and ST changes and time of delivery, related to perinatal outcome and metabolic acidosis in cases with adverse/complicated neonatal outcome *(continued)*

Case N:	Outcome	Metabolic acidosis	Interval between onset of changes and delivery (min)				Comments
			Intermediary CTG	Abnormal CTG	Preterminal CTG	ST events	
MAC 003	SCBU	No		722	233	177	7 episodes of biphasic ST in association with baseline FHR of 160 bpm with reduced reactivity and variability, no data last 33 minutes, maternal fever.
MAC 433	SCBU	No	56				Shoulder dystocia, normal ST, no data last 30 minutes

Specific FHR patterns associated with ST changes in cases of complicated adverse neonatal outcome:

The FHR patterns noticed above were also recorded in the group with cord artery metabolic acidosis but normal neonatal outcome. Among those cases, 1 case with late decelerations was recorded.

The US version of clinical guidelines has been found to accurately identify cases of metabolic acidosis and/or complicated/adverse neonatal outcome.

CTG, cardiotocography; *FHR*, fetal heart rate.

Amer-Wåhlin. Fetal electrocardiography ST analysis for intrapartum monitoring. Am J Obstet Gynecol 2019.

1. Borderline tachycardia with loss of variability and reactivity: 3 cases.
2. Tachycardia >170 bpm: 3 cases.
3. Episodes of persistent/prolonged bradycardia >2 minutes' duration: 11 cases.
4. Variable decelerations of >60 seconds' duration: 7 cases.
5. Variable decelerations of >60 bpm loss: 2 cases.
6. Loss of variability and reactivity (preterminal): 3 cases.