Combined cardiotocographic and ST event analysis: A review

Isis Amer-Wahlin, MD, PhD\textsuperscript{a}, Anneke Kwee, MD, PhD\textsuperscript{b,*}

\textsuperscript{a} Obstetrician, Medical Management Center and Department of Women and Child Health, ALB Q2:7, Karolinska Institute, Stockholm, Sweden
\textsuperscript{b} Obstetrician, Department of Obstetrics and Gynecology, University Medical Center Utrecht, Utrecht, The Netherlands

Keywords: ST-analysis, fetal ECG, STAN, cardiotocography, intrapartum fetal monitoring

ST-analysis of the fetal electrocardiogram (ECG) (STAN\textsuperscript{®}) combined with cardiotocography (CTG) for intrapartum fetal monitoring has been developed following many years of animal research. Changes in the ST-segment of the fetal ECG correlated with fetal hypoxia occurring during labor. In 1993 the first randomized controlled trial (RCT), comparing CTG with CTG + ST-analysis was published. STAN\textsuperscript{®} was introduced for daily practice in 2000. To date, six RCTs have been performed, out of which five have been published. Furthermore, there are six published meta-analyses. The meta-analyses showed that CTG + ST-analysis reduced the risks of vaginal operative delivery by about 10\% and fetal blood sampling by 40\%. There are conflicting results regarding the effect on metabolic acidosis, much because of controversies about which RCTs should be included in a meta-analysis, and because of differences in methodology, execution and quality of the meta-analyses. Several cohort studies have been published, some showing significant decrease of metabolic acidosis after the introduction of ST-analysis.

In this review, we discuss not only the scientific evidence from the RCTs and meta-analyses, but also the limitations of these studies. In conclusion, ST-analysis is effective in reducing operative vaginal deliveries and fetal blood sampling but the effect on neonatal metabolic acidosis is still under debate. Further research is needed to determine the place of ST-analysis in the labor ward for daily practice.

© 2015 Elsevier Ltd. All rights reserved.
Introduction

Electronic fetal heart rate (FHR) monitoring is a widely used method for assessing fetal status during labor. It aims to enable clinicians to identify hypoxic fetuses at risk for deterioration and provide prerequisites for a decision to intervene, and to deliver either vaginally or by cesarean section, thereby avoiding neonatal and long-term injury due to intrapartum asphyxia. Although little evidence exists regarding its efficacy, monitoring through cardiotocography (CTG) continues to be the method of choice in modern labor and delivery units in developed countries [1]. Despite this, there has been no significant reduction in the incidence of long-term neurologic morbidity (including cerebral palsy) and research has been stimulated by the low true positive predictive value of CTG for metabolic acidosis, which often results in unnecessary interventions and a significant increase in the cesarean section rate during the last 40 years due to concerns during labor. The addition of fetal blood sampling (FBS) is believed to hamper this effect; however, systematic reviews report no evidence of benefit in reducing the operative interventions [2]. Furthermore, FBS has also been shown to have a poor positive predictive value for intrapartum hypoxia [3]. This is probably due to the fact that performance of FBS requires expertise, is invasive, and must be repeated with persisting CTG abnormalities, and thus is often not performed when indicated [4].

Other tools for fetal surveillance, for example, fetal pulse oximetry, have not been successful maybe due to the well-known challenge lying in developing new and emerging technologies, related not only to the need to provide basic physiology data but also to meet requirements of data acquisition, signal processing, and data presentation [5]. Furthermore, any method in the fetal monitoring area requires understanding and compliance to clinical guidelines as well as a positive attitude toward changes of practice, a truly challenging perspective in such a medico-legally loaded field. Other important aspects related to evaluating the effect of medical technology are the choice of outcome parameters, the study design, the clinical setting of a trial, the ownership of the technique, as well as financial support available for its development. Moreover necessary clinical trials are expected to meet evidentiary standards that were never applied to existing technologies.

ST-analysis of the fetal electrocardiogram (ECG; STAN®) was introduced in the labor wards in 2000, after many years of research, starting with experimental animal research. Early animal studies observed that changes in the ST-segment of the fetal ECG correlated with fetal hypoxia occurring during labor [6,7]. The ST analyzer (STAN® monitor; Neoventa Medical, Goteborg, Sweden) was developed to combine traditional CTG with automatic analysis of the ST-segment of the fetal ECG. Changes in the shape of the ST-segment are noted automatically and an ST event is generated for a significant ST-change (Fig. 1). Guidelines have been developed defining whether intervention is required according to changes occurring in the CTG in combination with ST-changes of the fetal ECG (Table 1) [8].

Protocols to guide the use of a medical device have to be assessed and approved by the company responsible for the technology as part of the CE-marking process and by the FDA before any use in the United States. This implies that a change of guideline is limited by regulations similarly to change of an indication in relation to drugs. The premarketing approval (PMA) process, required for any new methodology, is the most stringent process that requires full documentation, including basic pathophysiology, signal processing, data presentation, control of device design, production, and software and adequate clinical data to support its safe and efficient use. On the basis of all these data and years of efforts, marketing approval is granted based on specific indications and method of use, and is enforced by law. Thus, it is not just the availability of a specific technology that allows a clinician to apply it, but the limit up to which it is approved for use.

The technique

The STAN® concept is based on the association between changes of the ST-interval of the fetal ECG and the function of the fetal myocardium during hypoxia. The changes in fetal ECG associated with fetal distress are either an increase in T-wave amplitude, quantified by the ratio of T-wave amplitude to QRS-amplitude (T/QRS ratio), or a biphasic ST-segment. An increase in T-wave amplitude and subsequently
in T/QRS ratio has been associated with a catecholamine surge, 
activation of β-adrenoreceptors, myocardial glycogenolysis, and metabolic acidosis [6,7]. A biphase shape of the ST-segment is related to two situations: (1) it may occur when the fetal heart is exposed to acute hypoxic stress, whereby it has had no time to respond to hypoxia, or (2) when the fetal heart has a reduced capacity to respond, due to (chronic) stress situations and lack of or already used resources. Biphasic ST-changes of the fetal ECG have been associated with disturbances in heart muscle function, infection, or malformations.

The STAN® concept is based on a combined interpretation of CTG and ST changes [8]. The relevance of an ST-change depends on the visual assessment of the CTG that, according to the criteria of the International Federation of Gynecology and Obstetrics (FIGO), is classified as “normal,” “intermediary” (or “suspicious”), “abnormal” (or “pathological”), and “(pre)terminal” [9]. If a CTG is normal, any ST-change on the STAN® monitor can be ignored. When a CTG is (pre)terminal, immediate intervention is advised, irrespective of ST-changes as is the case where the fetus has been assessed to be nonreactive or not found to be in a steady state at onset of the recording. In case of an intermediary or abnormal CTG, the STAN® guidelines indicate intervention in relation to ST-changes. Depending on the clinical circumstances, the intervention could be delivery, FBS, or alleviation of a possible cause of fetal distress, for example, uterine hypertonus or maternal hypotension. The STAN® guidelines can be used from a gestational age of 36 weeks onward.

Randomized clinical trials

Five RCTs, which included 15,365 patients, have been performed since 1993, all comparing continuous CTG-only monitoring with continuous CTG monitoring with ST-analysis [10–14]. Details of the studies are shown in Table 2. All trials used FBS in both arms.

Westgate et al. published the first randomized controlled trial (RCT) on the effect of intrapartum fetal ECG [10]. In this trial of 2434 women monitored with CTG or CTG + ST-analysis during labor, absolute values of T/QRS and visually verified trends of increase in T/QRS were used rather than automatically detected T/QRS changes, as developed during later trials and described in the current STAN clinical guidelines. In this first trial, the CTG + ST arm of the study showed a 46% decrease in the rate of operative vaginal deliveries and a trend toward less metabolic acidosis. Given these encouraging results, a second RCT was performed with a further development of the technique, including computerized ST-analysis (Table 2) and using guidelines based on the experience [11]. A significant decrease in neonatal metabolic acidosis demonstrated in the second RCT [11] was not supported by
findings in the following three RCTs [12–14]. These trials were not all powered to obtain a significant result and differed in the way metabolic acidosis was calculated (base deficit in blood giving considerably higher rates of metabolic acidosis compared with base deficit in extracellular fluid). Furthermore, one of the trials used abnormal CTG as inclusion criteria, without checking for baseline fetal

### Table 1
STAN® clinical guidelines.

<table>
<thead>
<tr>
<th>Cardiotocographic classification</th>
<th>Baseline heart frequency</th>
<th>Variability Reactivity</th>
<th>Decelerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>110–150 beats/min</td>
<td>5–25 beats/min</td>
<td>Early decelerations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accelerations</td>
<td>Uncomplicated variable decelerations with a duration of &lt;60 s and a beat loss of &lt;60 beats/min</td>
</tr>
<tr>
<td>Intermediary</td>
<td>100–110 beats/min</td>
<td>&gt;25 beats/min without accelerations</td>
<td>Uncomplicated variable decelerations with a duration of &lt;60 s and a beat loss of &gt;60 beats/min</td>
</tr>
<tr>
<td></td>
<td>150–170 beats/min</td>
<td>&lt;5 beats/min for &gt;40 min</td>
<td></td>
</tr>
<tr>
<td>Abnormal</td>
<td>150–170 beats/min and reduced variability</td>
<td>&lt;5 beats/min for &gt;60 min</td>
<td>Repeated late decelerations</td>
</tr>
<tr>
<td></td>
<td>&gt;170 beats/min</td>
<td>Sinusoidal pattern</td>
<td>Complicated variable decelerations with a duration of &gt;60 s</td>
</tr>
<tr>
<td>Preterminal</td>
<td></td>
<td></td>
<td>Total lack of variability and reactiveness with or without decelerations or bradycardia</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Intermediary CTG</th>
<th>Abnormal CTG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episodic T/QRS rise (duration &lt; 10 min)</td>
<td>Increase &gt;0.15 from baseline</td>
<td>Increase &gt;0.10 from baseline</td>
</tr>
<tr>
<td>Baseline T/QRS rise (duration ≥ 10 min)</td>
<td>Increase &gt;0.10 from baseline</td>
<td>Increase &gt;0.05 from baseline</td>
</tr>
<tr>
<td>Biphasic ST (a component of the ST-segment below the baseline)</td>
<td>Continuous &gt;5 min or &gt;2 episodes of coupled biphasic ST type 2 or 3</td>
<td>Continuous &gt;2 min or &gt;1 episode of coupled biphasic ST type 2 or 3</td>
</tr>
</tbody>
</table>

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

a Combination of several intermediary observations will result in an abnormal CTG.

### Table 2
Overview of randomized clinical trials.

<table>
<thead>
<tr>
<th>Authors/Year</th>
<th>Number of obstetric units/country</th>
<th>N</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Westgate J et al. 1993 [10]</td>
<td>1/UK</td>
<td>2434</td>
<td>Trend to decrease in metabolic acidosis (OR 0.38, 95% CI: 0.13–1.07) Increase in operative vaginal delivery rate (by 46%; p &lt; 0.001)</td>
</tr>
<tr>
<td>Amer-Wåhlin I et al. 2001 [11]</td>
<td>3/Sweden</td>
<td>4966</td>
<td>Decrease in metabolic acidosis (by 53%; p = 0.02) Decrease in operative vaginal delivery rate (by 17%; p = 0.047)</td>
</tr>
<tr>
<td>Ojala K et al. 2006 [12]</td>
<td>1/Finland</td>
<td>1483</td>
<td>No difference in metabolic acidosis No difference in operative delivery rate</td>
</tr>
<tr>
<td>Vayssière C et al. 2007 [13]</td>
<td>2/France</td>
<td>799</td>
<td>No difference in fetal blood sampling (by 56%; p &lt; 0.001).</td>
</tr>
<tr>
<td>Westerhuis M et al. 2010 [14]</td>
<td>9/The Netherlands</td>
<td>5681</td>
<td>No decrease in UA metabolic acidosis in ECF Decrease in UA metabolic acidosis in blood by 39% (RR 0.63, 95% CI 0.42–0.94) No difference in operative deliveries No difference in low Apgar score and neonatal HIE Decrease in FBS by 48% (RR 0.52, 95% CI 0.46–0.59)</td>
</tr>
</tbody>
</table>

UA: umbilical artery; ECF: extracellular fluid; OR: odds ratio; 95% CI: 95% confidence interval; ODFD: operative delivery for fetal distress; FBS: fetal blood sampling; HIE: hypoxic–ischemic encephalopathy; RR: relative risk.

a incl. acidosis, metabolic acidosis, low Apgar score, neonatal morbidity, and neonatal death.
status, thereby, in fact, violating the STAN guidelines, which state that ST-analysis should be started as early in labor as possible, preferably with a normal CTG [13].

The results of the five published STAN trials are summarized in Table 2. As can be observed, their results are quite different.

A large National Institute of Child Health and Human Development (NICHD)-driven multicenter trial (NCT01131260) in the United States presented its results recently at the Society for Maternal-Fetal Medicine (SMFM) 2015. There are some important differences between the study designs of the European trials and the US trial. In the FDA-approved US trial, a simplified approach with a 3-category CTG classification system was chosen, thus assessing nonreassuring FHR (NRFHR grade 1) in one category instead of separating them into intermediary and abnormal as in the European 4-category guidelines. The aim was to simplify the interpretation of CTG + ST changes and reduce the risk for ambiguous data interpretation. However, major changes occurred in the US user guidelines affecting the decision to act based on ST-waveform changes as an adjunct to CTG analysis. The most obvious is that a 60’ rule has been included, stating that in case of an NRFHR for >60’ but no ST event, direct physician assessment of fetal state was required with intrauterine resuscitation, and if no improvement was observed, expeditious delivery was performed.

The statement in the US study protocol that could cause ST-analysis to be completely ignored was a paragraph stating the following: 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 FHR-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. As a consequence, one would assume unnecessary interventions, especially in 2nd stage with the risk of neonates being affected due to emergency operative procedures. Furthermore, the following statement was also added in the US study protocol:

In the presence of maternal fever and related infection, the fetus may have a blunted or no response to hypoxia, and ST events may fail to appear. Therefore, if the maternal temperature reaches 38.0°C (100.4°F) or greater, management should be related to the FHR and the clinical situation.

Thus, the USRCT study protocol was substantially altered from the protocol approved by the FDA, and the guidelines at some decisive points different from those used in the previous STAN RCTs, making comparisons problematic.

Meta-analysis of the RCTs

Six meta-analyses [15–20], including these five RCTs, have been performed, of which one is an individual patient data meta-analysis (IPDMA) [19] and one is a “correction” of errors performed in the first five meta-analyses [20]. Details are shown in Table 3. The IPDMA offers numerous statistical and clinical advantages over an aggregate data meta-analysis as it increases the power to detect differential

Table 3
Overview of meta-analysis.

<table>
<thead>
<tr>
<th>References of RCTs included</th>
<th>RR (95% CI) metabolic acidosis</th>
<th>RR (95% CI) Operative delivery</th>
<th>RR (95% CI) Fetal blood sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neilson [15]</td>
<td>0.78 (0.44–1.37)</td>
<td>0.89 (0.81–0.98)</td>
<td>0.61 (0.41–0.91)</td>
</tr>
<tr>
<td>Becker [16]</td>
<td>0.72 (0.43–1.19)</td>
<td>0.99 (0.91–1.08)</td>
<td>0.59 (0.44–0.79)</td>
</tr>
<tr>
<td>Potti [17]</td>
<td>0.80 (0.44–1.47)</td>
<td>1.03 (0.67–1.21)</td>
<td>0.59 (0.44–0.79)</td>
</tr>
<tr>
<td>Salmelin [18]</td>
<td>0.96 (0.49–1.88)</td>
<td>0.89 (0.83–0.97)</td>
<td>NM</td>
</tr>
<tr>
<td>Schuit [19]</td>
<td>0.76 (0.53–1.10)</td>
<td>0.89 (0.83–0.97)</td>
<td>0.55 (0.4–0.76)</td>
</tr>
<tr>
<td>Olofsson [20]</td>
<td>0.61 (0.41–0.91)</td>
<td>0.89 (0.81–0.95)</td>
<td>0.64 (0.47–0.88)</td>
</tr>
</tbody>
</table>

| a | Mixture of BDblood and BDecf. |
| b | Total operative vaginal deliveries. |
| c | Total cesarean section. |
| d | Total operative deliveries for fetal distress. |
treatment effects across individuals in RCTs. The IPDMA showed that CTG + ST-analysis of the fetal ECG reduces the risk of metabolic acidosis in the extracellular fluid (RR 0.76; CI 95% 0.53–1.10) compared with CTG alone, although the reduction was not statistically significant. The numbers of FBS (RR 0.49; CI 95% 0.44–0.55) and instrumental vaginal deliveries (RR 0.90; CI 95% 0.83–0.99) were significantly reduced by CTG + ST-analysis. Furthermore, CTG + ST-analysis reduced the incidence of metabolic acidosis calculated in blood, arterial pH < 7.15, arterial pH < 7.05, arterial pH < 7.00, NICU admissions, hypoxic–ischemic encephalopathy, need for intubation, seizures, and a composite of adverse perinatal outcome, although not statistically significant. The cesarean section rate was comparable in both groups (RR 0.99; CI 95% 0.91–1.09). For the other secondary outcomes, no substantial differences were found between CTG + ST-analysis and CTG alone. In addition, CTG + STAN reduced the incidence of neonatal intensive care admission for infants born after 41 weeks of gestation (RR 0.61; CI 95% 0.39–0.95).

However, in order to perform a reliable meta-analysis, the included studies should address the same research question, be of comparable quality regarding selection bias, attrition rates, and confounding variables, and include comparable populations. Unfortunately, the different STAN trials have considerable methodological discrepancies [65], thereby reducing the strength of a meta-analysis. As such, for example, the RCTs included a wide spectrum of patients, from mixed-risk patients requiring amniotomy to high-risk patients with suspicious CTG recordings. One trial restricted the use of FBS in the CTG + STAN group to strict criteria that included intermediary or abnormal CTG at start of the registration, abnormal CTG for more than 60 min in the absence of ST events in the 1st stage of labor, and poor ECG quality in the presence of an intermediary or abnormal CTG trace.

In the latest meta-analysis [20], the methodology, execution and quality of the first five meta-analyses were scrutinized. These first meta-analyses contained errors, either created in handling of the original RCT data or from a failure to recognize essential differences in methodology among the RCTs. After correction of the uncovered errors and exclusion of the RCT using deviant STAN guidelines for enrollment, the latest meta-analysis showed that CTG + ST monitoring significantly reduces the need for FBS (reduction of 36%) and operative delivery (reduction of 7%), and reduces the neonatal metabolic acidosis rate (reduction of 39%) [20].

Cohort studies

Since the introduction of STAN, several cohort studies are published [4,21–37], as summarized in Table 4. Recent observational studies investigating the effects of long-term use of ST-analysis have shown a reduction in the incidence of metabolic acidosis (base deficit in extracellular fluid (BDect)) over time.

Nevertheless, it should be kept in mind that, as with all new medical technologies requiring user input, a significant learning curve exists. Clinical success may be highly dependent on the skill of the clinician. As skills evolve, observed effectiveness of the technology may improve. Furthermore, the incidence of adverse events may decrease, improving the overall benefit–risk relationship of the procedure. The heterogeneity in clinical safety and effectiveness related to the combined source of variation present in the device–operator interaction has been recognized since long [38].

As such, results of STAN may improve over time as clinicians become more comfortable in integrating with the tools. This is reflected in two observational studies from Sweden and Norway, describing increasing STAN use over time and extremely low metabolic acidosis rates [31,34]. Undoubtedly, such positive development requires an adequate organization and continuous teaching and learning.

ST-analysis in particular situations

Immature or anatomically abnormal hearts

ST-analysis in labor is based on the presumption that the fetal ECG waveform in the examined fetus is normal and that the fetal myocardium will react to hypoxia by switching from an aerobic to anaerobic metabolism with glycogenolysis. As such, this precludes analysis in fetuses with a high
## Table 4
Observational studies reporting the clinical use of cardiotocography and ST-analysis of fetal ECG in labor.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Type of study</th>
<th>Number of obstetric units/country</th>
<th>N</th>
<th>Main results and conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luzietti R et al. 1999 [21]</td>
<td>Prospective</td>
<td>7/European Union</td>
<td>320</td>
<td>27 cases had ST changes including all six cases of intrapartum asphyxia.</td>
</tr>
<tr>
<td>Amer-Wåhlin I et al. 2002 [22]</td>
<td>Prospective</td>
<td>12/Norway, Sweden</td>
<td>573</td>
<td>ST-analysis plus CTG had higher positive predictive value for metabolic acidosis than CTG alone.</td>
</tr>
<tr>
<td>Kwee A et al. 2004 [23]</td>
<td>Prospective</td>
<td>1/The Netherlands</td>
<td>637</td>
<td>ST changes in all five cases with severe metabolic acidosis and in 46% of cases with mild metabolic acidosis; CTG + ST more specific than CTG alone.</td>
</tr>
<tr>
<td>Ross MG et al. 2004 [24]</td>
<td>Retrospective</td>
<td>7 clinicians/USA</td>
<td>51</td>
<td>ST-analysis in addition to FHR monitoring improves consistency in decision making and timing of obstetric interventions.</td>
</tr>
<tr>
<td>Devoe LD et al. 2006 [25]</td>
<td>Prospective</td>
<td>6/USA</td>
<td>530</td>
<td>The negative predictive value for avoiding intervention in fetuses with a nonreassuring FHR pattern was 95%.</td>
</tr>
<tr>
<td>Norén H et al. 2006 [26]</td>
<td>Prospective</td>
<td>2/Sweden</td>
<td>4830</td>
<td>Decrease in metabolic acidosis (p &lt; 0.05). Decrease in overall and emergency C/S rates (OR 0.18; 95% CI 0.07–0.49). Low metabolic acidosis rate (0.5%).</td>
</tr>
<tr>
<td>Welin AK et al. 2007 [27]</td>
<td>Retrospective</td>
<td>1/Sweden</td>
<td>1875</td>
<td>Decrease in instrumental deliveries for fetal distress (p &lt; 0.025), admissions to NICU (p &lt; 0.01) and in low Apgar scores (p &lt; 0.001). No difference in C/S rate for fetal distress.</td>
</tr>
<tr>
<td>Kale A et al. 2008 [28]</td>
<td>Prospective</td>
<td>1/Singapore</td>
<td>474</td>
<td>Decrease in instrumental deliveries for fetal distress (p &lt; 0.025), admissions to NICU (p &lt; 0.01) and in low Apgar scores (p &lt; 0.001). No difference in C/S rate for fetal distress.</td>
</tr>
<tr>
<td>Massoud M et al. 2007 [29]</td>
<td>Prospective</td>
<td>1/France</td>
<td>1889</td>
<td>Low metabolic acidosis rate (0.38%); trend to decrease in ODFD.</td>
</tr>
<tr>
<td>Doria V et al. 2007 [30]</td>
<td>Prospective</td>
<td>1/UK</td>
<td>1502</td>
<td>Metabolic acidosis occurred in 2.8%, and 70% were identified by STAN. Retrospective analysis of the 14 cases with encephalopathy monitored by STAN showed that the guidelines for intervention were not followed.</td>
</tr>
<tr>
<td>Norén H et al. 2007 [4]</td>
<td>Retrospective</td>
<td>8/EU</td>
<td>911 cases with FBS</td>
<td>Number of operative deliveries decreased without an increase in the number of newborns with metabolic acidosis.</td>
</tr>
<tr>
<td>Palmgren Colov NS 2007 [31]</td>
<td>Prospective</td>
<td>1/Denmark</td>
<td>1168</td>
<td>ST-analysis + CTG as effective as fetal blood sampling.</td>
</tr>
<tr>
<td>Melin M et al. 2008 [32]</td>
<td>Retrospective</td>
<td>1/Sweden</td>
<td>506</td>
<td>An ST event occurred in 79% of severe and 75% of moderate metabolic acidemia cases, as well as in 50% of controls. Two of three cases with severe and &lt;50% of cases with moderate metabolic acidemia were preceded by ST events coinciding with CTG abnormalities.</td>
</tr>
<tr>
<td>Rzepka R et al. 2010 [33]</td>
<td>Prospective</td>
<td>1/Poland</td>
<td>83</td>
<td>ST-changes reflect metabolic acidosis better than CTG.</td>
</tr>
<tr>
<td>Norén H &amp; Carlsson A 2010 [34]</td>
<td>Prospective</td>
<td>1/Sweden</td>
<td>12,832</td>
<td>STAN in 26–69%</td>
</tr>
<tr>
<td>Ragupathy K et al. 2010 [35]</td>
<td>Prospective</td>
<td>1/UK</td>
<td>253</td>
<td>1.5% metabolic acidosis at birth, 100% preceded by ST events but guidelines not followed. High percentage of ST events in laboring women.</td>
</tr>
<tr>
<td>Doret et al. 2011 [36]</td>
<td>Prospective</td>
<td>1/France</td>
<td>3112</td>
<td>No FBS performed. A total of 14 cases with metabolic acidosis of which 11 not managed according to guidelines. CS rate fetal distress 9.5%. Study supports the use of ST analysis without FBS, but warns for guideline violations.</td>
</tr>
</tbody>
</table>
likelihood of abnormal ECG shapes, such as fetuses with cardiomyopathies and cardiac defects. Less evidently, the preterm myocardial metabolism is also different from the term metabolism because of a less-advanced cardiomyocyte maturation [39]. Our knowledge of the changes in the ECG waveform as a reaction to hypoxia in preterm fetuses is limited so far. ST-analysis is therefore restricted for use at term.

Maternal diabetes

Fetuses of mothers with pre-pregnant diabetes are at risk of hypertrophic cardiomyopathy, which is known to be associated with ECG changes postnatally. A case-control subanalysis of women with gestational diabetes \( n = 338 \) and pre-pregnant diabetes \( n = 75 \) participating in two STAN trials looked further into this topic and indeed showed that ST-segment depression is more common in fetuses of mothers with pre-pregnant diabetes (22% vs. 12% in controls), than with gestational diabetes [40]. No difference was found in the incidence of ST-elevation between the two types. Although the study was underpowered to assess the influence of these changes on the sensitivity and specificity of ST-analysis for the detection of hypoxia in utero, the altered physiology needs to be kept in mind when assessing the fetus of a diabetic mother. So far, maternal diabetes is not considered as an exclusion criterion for STAN.

Maternal fever in labor

Maternal fever, which is sometimes associated with fetal infection, could theoretically affect the fetal ECG waveform, by either direct effects on the myocardium or a higher incidence of hypoxia. A case-control subanalysis of two Dutch STAN studies however showed that, despite a higher occurrence of fetal infections in a cohort of mothers with intrapartum fever \( n = 142 \), the incidence of ST-changes was similar to controls [41]. Therefore, maternal fever should not be considered as a contraindication for STAN. One should nevertheless keep in mind that chorioamnionitis and fetal infection are known risk factors for adverse cerebral events, independent of hypoxia, and that the level of attention in relation to both CTG and ST-analysis in these situations should be increased.

Growth-restricted fetuses

The cardiac resources of a growth-restricted fetus, which is exposed to chronic stress, may already be exhausted, and the fetus may not be able to respond to additional stress of labor with the usual increase in T/QRS ratio. This was confirmed in a guinea pig model of intrauterine growth restriction (IUGR) [42]. Fetuses with IUGR however still respond to additional stress, not by ST-elevation but by ST-depression and biphasic events, and can therefore still be monitored with the current STAN algorithms.

Epidural analgesia

Theoretically, epidural analgesia can have an influence on the ECG pattern, as a result of either hypotension-induced fetal hypoxia or medication-related ECG changes. The placental intervillous space is known to act as a site of drug deposit for bupivacaine as well as fentanyl, and effects of bupivacaine on the fetal ECG are described. Besides a significant increase in FHR, a significant change in

Table 4 (continued)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Type of study</th>
<th>Number of obstetric units/country</th>
<th>N</th>
<th>Main results and conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kessler et al. 2012 [37]</td>
<td>Prospective</td>
<td>1/Norway</td>
<td>6010</td>
<td>During the study period (4 years), metabolic acidosis rate decreased (1.4–0.3%) and cesarean section rate decreased (10.1–8.8%)</td>
</tr>
</tbody>
</table>

CTG: cardiotocography; FHR: fetal heart rate; C/S: cesarean section; OR: odds ratio; 95%CI: 95% confidence interval; NICU: neonatal intensive care unit; ODFD: operative delivery for fetal distress; FBS: fetal blood sampling.
T/QRS ratio was also observed. It is uncertain if there is a significant fetal metabolism or drug uptake by the fetus, despite the high rates of placental transfer. Becker et al. compared 72 women with and without epidural analgesia [43]. No significant difference was found between the two groups regarding the numbers of ST events, types of ST events, and significant ST events (intervention advised) or nonsignificant. The authors concluded that epidural analgesia has no effect on the numbers or types of ST events when using ST-analysis.

Fetus in breech position

One cohort study describes the use of ST-analysis in 433 deliveries with fetuses in breech position [44]. Compared with deliveries with fetuses in vertex position, a lower risk of T/QRS baseline rise during labor (OR 0.7; CI 95% 1.6—5.9) was found, with comparable and acceptable neonatal outcomes. There was a higher risk of intervention due to preterminal CTG patterns. The authors concluded that ST-analysis can be used for vaginal delivery of a fetus in breech presentation.

The ECG pattern varies with the fetal presentation, even after adjusting for gestational age and birthweight. A possible explanation for this finding is that ST-interval changes evolve not only as a result of hypoxia, but also as a result of fetal arousal caused by physiological sympathetic stimulation during labor. The effect of a given external stimulus may depend on fetal presentation: the response to vibroacoustic stimulation close to term has been found to be lower in breech than in cephalic presentations [45]. Furthermore, the sensitivity of somatosensory perceptions in humans is higher in the face than in the lower extremities and buttocks [46]. As a consequence, the mechanical forces affecting the presenting part may result in higher levels of fetal arousal when applied to the fetal skull, eyes, and face than to the feet and buttocks, which could explain the different patterns of ST-interval changes between breech and cephalic presentations.

Cost-effectiveness

Reports of two slightly different cost-effectiveness analyses have been published, based on the results of two of the larger randomized trials [47,48]. In essence, both studies balanced the cost of STAN equipment and operator training in ST-analysis against the costs of FBS, vacuum extraction, and metabolic acidosis. Neither of the studies considered any cost for operator training in CTG interpretation. The first study, based on the Swedish trial, found ST-analysis to be the most cost-effective procedure with a 56-euro lower cost per patient and a gain of 0.005 quality-adjusted life years [47]. The second study, based on the Dutch trial, estimated an additional cost of 29 euro per labor monitored with ST-analysis, which would result in a cost of 9667 euro to prevent one case of metabolic acidosis [48]. Both studies, therefore, conclude that ST-analysis in labor is probably cost-effective from a population perspective. The results of these analyses are however based on the results of individual randomized trials, which showed improvements in neonatal outcomes and a decrease in metabolic acidosis. The latter was not confirmed in all the meta-analyses, and the question therefore remains whether these cost-effective analyses can be extended to daily practice.

Studying medical technology

Any systematic review of evidence needs to take into account the quality of the evidence. Any study, whether randomized or observational, may have flaws such as problems in methods of recruiting patients, in the clinical setting, or in the delivery of the treatment that can cast doubt on the generalizability of the results. Heterogeneity and publication bias are relevant to all comparisons of evidence from both randomized, controlled studies and observational studies. A current consensus exists in medicine regarding the hierarchy of study designs in clinical research. Randomized, controlled studies are believed to represent the gold standard applying to all clinical situations; however, this has been questioned in relation to studies of medical technology as the impact is at least partially related to the technical complexity of the device combined with operator training and decision-making [49]. In a randomized double-blind drug trial, the clinical effect observed is imparted by the molecular structure of the compound. Change the compound, even slightly, and assumptions about replication of clinical
effect are made at the clinician’s risk. The evidence generation for how technology improves outcome is an environment quite different from that for drug evaluations. Clinical trials of implementing technology in the obstetric setting are conducted in a relatively unique environment defined by the clinicians acting in that environment.

Technical trials will rely on information provided on a screen on which the clinician will act, and thus are methodologically challenging; the fundamental concepts of randomization, blinding, and allocation concealment are not easily implementable. Further, unblinded randomization between a technology and another, such as in all the described RCTs, poses questions regarding the credibility as so much of the outcome will depend on the clinician’s judgment [50].

The skills and experience of the clinician are trial limitations deriving from factors external to the technology itself. The assignment of “treatment” when monitoring labor is left to the judgment of the clinician. The clinician makes the choice of therapy based on preferences, training, and institutions rather than strictly based on available information, disease-related factors, or the facilities available.

There is substantial evidence in the literature that willingness of the clinician to intervene, rather than its appropriateness for the patient, may explain substantial variation in practice [51,52]. In obstetrics, the factors influencing physicians’ decision to perform a cesarean section for fetal distress have been studied, and practice style and personal attitude have been reported to be of great importance [53,54].

In a study on management of twin pregnancies, despite the diagnostic progresses in the field of multifetal pregnancies, management remained to a large extent centre dependent, arbitrary, and influenced by anecdotal experience and medicolegal issues [55]. Thus, clinical decision making is a complex arena that does not lend itself to easy interpretations and there seems to be an agreement that implementation of changes is difficult even when the evidence clearly supports those changes [56].

In literature, little is found related to practice and consequences of clinical research. During the Swedish RCT [11], an interim analysis performed after 1800 patients which revealed several protocol violations. A protocol violation was defined as a situation where intervention was performed despite a reassuring ST-analysis or where intervention was not performed when there was an indication to intervene, in conflict with the study protocol and clinical guidelines. After completion of the trial, rates of metabolic acidosis and operative interventions in both study arms before and after the interim analysis were compared. No difference in the occurrence of inadequate recordings was found throughout the study, suggesting “technical” knowledge was satisfactory. However, adherence to the clinical guidelines was a challenge, as shown by the protocol violations revealed at the interim analysis, and a statistically significant change over time in the rate of operative deliveries for fetal distress \( p = 0.02 \). Further a decreased number of fetal blood samples \( p = 0.001 \) was observed over time, suggesting a gradual improvement of performance throughout the study. The impact of training and experience is illustrated by data from the IPDMA: when analyzing the total data there were no significant differences between the CTG + ST group and the CTG only group regarding neonatal metabolic acidosis, low cord artery pH, admission for neonatal intensive care, or a composite adverse outcome parameter, but when data from only the second halves of the RCTs were analyzed, there were significant reductions in all parameters by 24–50% in benefit of CTG + ST monitoring [57]. This can be seen as an expression of barriers to behavioral change necessary for acceptance existing during an otherwise accepted trial. Such barriers have been classified earlier into three categories: environmental, prevailing opinion, and personal characteristics [58]. Of the environmental barriers, in modern obstetrics, the patient expectations and perception of liability can be very dominant. Adverse outcome of pregnancy and delivery is a very rare but loaded phenomenon. The prevailing opinion barriers are usually based on the medical training of the staff. Furthermore, the existing standards of practice and autocratic medicine principles used locally can be hostile to changes and even to a trial protocol. Finally, the personal characteristic barriers, expressed as individual practice style, can be very prominent among obstetricians [53]. The situation in the labor ward exposes clinicians to clinical uncertainty. A feeling of discomfort and a compulsion to act, sometimes in conflict with the guidelines, can emerge. All the aforementioned barriers were observed in the Swedish RCT, and overall, there were differences in the way the obstetricians and midwives approached the new method according to age, sex, and personal characteristics [59].

The attitude of staff using STAN and the need of training, practice, motivation, and desire to use new technologies have been published previously [60,61]. The authors report the various strategies needed to obtain adherence to protocol, such as discussions and reviewing recent tracings.
Thus personal characteristics of individuals participating in RCTs and the social influences among the staff can be of paramount importance. Naturally, differences exist in individual receptivity, perception of risks, and acceptance of uncertainty. Factors that might have had an effect on the performance within the trial, as previously described from another RCT in the field of obstetrics [62].

An RCT lasting for several years necessitate integration of all members of staff in the clinical teamwork of the labor ward to promote the social diffusion of the new practice. In relation to fetal monitoring trials, the staff members were exposed equally to the education, pretrial training, ongoing training within the RCT, and feedback from own cases, in audit with group discussions or individually. Coleman et al. describe that clinicians will be influenced more by what the colleagues say and do in uncertain situations than in clear-cut situations [63]. Uncertain situations are common vis-à-vis a patient in labor.

In the Swedish RCT, it was obvious that the increasing personal experience resulted in increased confidence in using the STAN technology and better acceptance of the guidelines. The subsequent detailed comparative analysis of the CTG and ST-recordings showed that the ST-recordings gave a better inter- and intraobserver reproducibility, and better specificity regarding the pregnancy outcome than CTG alone [64].

The study also showed that a proper performance within the RCT, as well as changing professional behavior in clinical practice, cannot be brought about by relatively passive methods of disseminating new knowledge, for example, by individual studies of new guidelines. Permanent active involvement of the principal investigators was necessary to ensure a smooth course of the RCT. In the complex clinical world, the audit performed of cases from the study was a good example of “single-loop” learning strategy, enabling detection and correction of errors and increasing the level of understanding and compliance with the guidelines. For implementation of new methods into clinical practice, “double-loop” learning strategy might be advantageous because the new guidelines are presented in a new pattern of interactions and values where the goals of the whole organization are being redefined, and the actual norms and policies are being questioned [65].

Thus, analyses of two of the performed RCTs indicate that conventional training did not have the expected profound effect. It was rather the audit of cases, the ensuing discussions, and the staff’s increasing personal experience with the use of the method that led to improved acceptance of clinical guidelines during the RCT. For successful trials on technical innovations, it is probably necessary to guarantee a certain level of scientific knowledge of the recipients in order to achieve optimal effect of educational activities. In addition, barriers of implementation must be overcome by increasing personal familiarity of all participants with the new method and changing the existing culture, practice, and routines.

The popular belief that only RCTs produce trustworthy results could imply a disservice to patient care and clinical investigation. Observational studies have their own unique challenges related to the interaction between technology, clinician, and the patient. Technical complexity, information processing, and clinical details are important when evaluating outcome, and missing information about these variables threaten the quality of any study. Observational studies have one crucial deficiency: the design is not an experimental one. Each patient’s treatment is deliberately chosen and not randomly assigned, and unavoidable risks of selection bias and systematic differences in outcomes exist that are not due to the treatment itself. However, in the light of the described risk of bias in any trial related to the clinicians’ performance, a strategy to overcome the latter limitation is to exploit the natural variation in care in quality improvement research, such as observational studies describing routine work [34]. Such studies are often performed in one center with the benefit of similar preferences, training, and institution, minimizing the probability of variation over time once a technology, such as EFM, has been implemented into practice.

Conclusion

ST-analysis has been extensively studied in both RCTs and observational studies. The results are conflicting. Most meta-analyses show that additional use of ST-analysis for intrapartum monitoring does not reduce the incidence of metabolic acidosis, but does reduce the incidence of operative vaginal deliveries and the need for FBS. However, criticism related to the published evidence in terms of
interpretation of studies introduces a doubt if evidence in the case of fetal monitoring will ever be black or white before long-term outcome is known [20,66].

The definite results of the US trial and the impact of these results on labor wards using the evidence-based STAN guidelines (as was not the case in the US trial) and additional FBS are still unknown. There are promising cohort studies showing a decrease of metabolic acidosis over time. Although it is difficult to definitely prove that this is caused by the use of ST-analysis, cohort studies are probably reflecting the situation over time in a labor ward in a better way than that of an RCT.

Conflicts of interest

There are no conflicts of interest.

Practice points

- Changes in the ST-segment of the fetal ECG are associated with fetal hypoxia.
- The STAN concept is based on a combined interpretation of CTG and ST-changes.
- Use of CTG+ST-analysis reduces the risk of vaginal operative delivery and the number of fetal blood samples.
- There are particular situations, such as anatomically abnormal heart, maternal diabetes, maternal fever, or growth-restricted fetuses, in which the development of ST-changes can be different from that of fetuses without these complications.
- The context of the clinical situation always has to be kept in mind when interpreting CTG and ST-analysis.
- There are 5 published RCT’s comparing CTG+ST analysis with CTG only, in all trials FBS was possible. The latest 4 RCTs used common clinical guidelines.
- The results related to metabolic acidosis are conflicting. Calculation of base deficit differed between the meta-analysis performed.

Research agenda

- A new meta-analysis after publication of the US trial.
- Thorough discussion and evaluation of the results of the US trial including audit of adverse cases.
- Studies of effectiveness of ST-analysis in specific groups, such as fever during labor, growth-restricted fetus, and premature fetus.
- Research related to the understanding of successful implementation strategies.
- Research to further improve the technique of ST-analysis.

References


