Received: 25 September 2023 | Revised: 25 October 2023 | Accepted: 23 November 2023

DOI: 10.1111/aogs.14752

AOGS Acta Obstetricia et Gynecologica Scandinavica

SYSTEMATIC REVIEW

ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: An updated systematic review and meta-analysis of randomized trials

Ellen Blix[1] | Kjetil Gundro Brurberg[2,3] | Eirik Reierth[4] | Liv Merete Reinar[2] | Pål Øian[5]

[1]Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

[2]The Norwegian Institute of Public Health, Oslo, Norway

[3]Center for Evidence Based Practice, Western Norway University of Applied Sciences, Bergen, Norway

[4]Science and Health Library, University Library, UiT The Arctic University of Norway, Tromsø, Norway

[5]Department of Obstetrics and Gynecology, University Hospital of North Norway, Tromsø, Norway

Correspondence
Eirik Reierth, Science and Health Library, University Library, UiT The Arctic University of Norway, P.O. Box 6050 Langnes, NO-9037 Tromsø, Norway.
Email: eirik.reierth@uit.no

Abstract

Introduction: ST waveform analysis (STAN) was introduced as an adjunct to cardiotocography (CTG) to improve neonatal and maternal outcomes. The aim of the present study was to quantify the efficacy of STAN vs CTG and assess the quality of the evidence using GRADE.

Material and methods: We performed systematic literature searches to identify randomized controlled trials and assessed included studies for risk of bias. We performed meta-analyses, calculating pooled risk ratio (RR) or Peto odds ratio (OR). We also performed post hoc trial sequential analyses for selected outcomes to assess the risk of false-positive results and the need for additional studies.

Results: Nine randomized controlled trials including 28 729 women were included in the meta-analysis. There were no differences between the groups in operative deliveries for fetal distress (10.9 vs 11.1%; RR 0.96; 95% confidence interval [CI] 0.82–1.11). STAN was associated with a significantly lower rate of metabolic acidosis (0.45% vs 0.68%; Peto OR 0.66; 95% CI 0.48–0.90). Accordingly, 441 women need to be monitored with STAN instead of CTG alone to prevent one case of metabolic acidosis. Women allocated to STAN had a reduced risk of fetal blood sampling compared with women allocated to conventional CTG monitoring (12.5% vs 19.6%; RR 0.62; 95% CI 0.49–0.80). The quality of the evidence was high to moderate.

Conclusions: Absolute effects of STAN were minor and the clinical significance of the observed reduction in metabolic acidosis is questioned. There is insufficient evidence to state that STAN as an adjunct to CTG leads to important clinical benefits compared with CTG alone.

KEYWORDS
cardiotocography, fetal electrocardiography, intrapartum fetal monitoring, living systematic review, meta-analysis, ST waveform analysis

Abbreviations: CTG, cardiotocography; GRADE, Grading of Recommendations, Assessment, Development and Evaluation; MeSH, Medical Subject Headings; NICU, neonatal intensive care unit; RCT, randomized controlled trial; STAN, ST waveform analysis; TSA, trial sequential analyses.

1 | INTRODUCTION

The aim of fetal monitoring is to identify fetuses at risk of neonatal and long-term injury attributable to asphyxia and enable timely interventions to prevent cases of fetal damage.

Cardiotocography (CTG) was introduced in the 1960s and assumed to prevent fetal asphyxia, and soon became widely used in clinical practice. The use of CTG has been associated with a decrease in neonatal seizures after prolonged labor but not with improved long-term outcomes in the child. It has also been associated with an increase in cesarean sections and assisted vaginal deliveries.[1] The CTG method has limitations such as low specificity, high false-positive rates and high inter-rater variability; therefore, a method with better diagnostic accuracy is needed to identify truly hypoxic fetuses.

The ST waveform analysis (STAN) method was introduced after extensive experimental research in Sweden.[2] Following a lack of fetal oxygen, anaerobic metabolism will produce changes in the fetal ECG. The method can be used after rupture of membranes in single pregnancies after 36 weeks' gestation. A scalp electrode is necessary for monitoring.

In 2015 and 2016, three systematic reviews and meta-analyses comparing CTG + STAN vs CTG were published.[3–5] The three meta-analyses included the same six randomized controlled trials[6–13] and arrived at the same conclusions: that the absolute effect of CTG + STAN on neonatal outcomes was minor compared with CTG alone. There was a reduction in babies born with metabolic acidosis in cord blood in women allocated to the CTG + STAN group; the relative risk reduction was 36% and the absolute risk reduction 0.25%. The difference was statistically significant in only one of the three meta-analyses (0.45% vs 0.7%, Peto OR 0.64, 95% CI 0.46–0.88).[3] There were no differences in other neonatal outcomes, such as Apgar scores, neonatal seizures or encephalopathy, or transfer to a neonatal intensive care unit. Women randomized to the STAN group had a lower risk of having fetal blood samples taken during labor, and for an assisted vaginal delivery for all indications compared with women randomized to the CTG group.[3–5] There were no differences in maternal outcomes such as cesarean section rates or assisted vaginal deliveries for fetal distress.[3–5]

A newer systematic review and network meta-analysis evaluated the effectiveness of different types of fetal monitoring.[14] It reported that intermittent auscultation reduced the risk of emergency cesarean sections without compromising neonatal outcomes compared with other methods, except when compared with CTG in combination with STAN and fetal blood sampling. However, in two of the seven studies included in the CTG + STAN group in the network meta-analysis, the fetal ECG analyses were of the PR segment and not the ST segment. Therefore, the results should be interpreted with caution.

We aimed to update our previous systematic review[3] to quantify the efficacy of the STAN method as an adjunct to conventional CTG compared with CTG alone. In addition to conventional quality assessments of the included studies, we used the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) to assess the overall quality of evidence.[15] We performed trial sequential analyses (TSA) for selected outcomes to assess the risk of false–positive results, futility and the need for additional trials.[16]

Key message

It is unclear whether ST waveform analysis is better for labor surveillance than conventional CTG. Evidence is insufficient to state that STAN as an adjunct to CTG leads to important clinical benefits compared with CTG alone.

2 | MATERIAL AND METHODS

We updated our previous systematic review.[3] The protocol is published in the PROSPERO international prospective register of systematic reviews, registration no. CRD42015023563.

We repeated our previous literature searches, following the same search strategy, in the following databases: Ovid MEDLINE® (In-Process & Other Non-Indexed Citations, Ovid MEDLINE®, Daily, Ovid MEDLINE® and Ovid OLDMEDLINE®), EMBASE Classic+ (EMBASE (Ovid), The Web of Science® (Thomson Reuters), The Cochrane Library (Wiley) and CINAHL Plus (EBSCO*host*). The searches were performed on October 31 2022, with the limitation 2015–2022 (current), and new searches were performed on June 22, 2023. The searches are described in detail in Appendix S1.

2.1 | Study selection and data extraction procedures

The citations identified by the systematic searches were screened and potentially eligible studies were obtained in full text for further assessment. Two reviewers (EB, PØ) assessed eligibility of the studies independently. Persisting disagreements were resolved by consulting a third reviewer (LMR). The selection criteria were:

- Population: 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;
- Intervention: CTG plus STAN;
- Comparator: CTG alone;
- Primary outcomes: operative deliveries for fetal distress, metabolic acidosis in the newborn (pH <7.05 and $BD_{(ecf)}$ >12 mmol/L in umbilical artery). Secondary outcomes: neonatal and perinatal

death, neonatal seizures, neonatal encephalopathy, transfers to the neonatal intensive care unit (NICU), fetal blood sampling, cesarean sections, operative vaginal delivery, Apgar score <7 at 5 minutes and a composite (ie intrapartum death, neonatal death, Apgar score <4 at 5 minutes, neonatal seizures, metabolic acidosis, intubation at delivery for ventilation or neonatal encephalopathy);
- Study design: randomized controlled trial (RCT).

Two of the reviewers (EB, PØ) extracted data from each study independently, using a predesigned form.

2.2 | Assessments and synthesis

All studies meeting the selection criteria were critically appraised using the Risk of Bias tool developed and recommended by the Cochrane Collaboration.[17] Two reviewers (EB, LMR) performed the risk of bias assessments independently. Persisting disagreements were resolved by consulting a third reviewer (KGB).

Numbers of mothers or infants with the outcome of interest were extracted from the included studies and entered into a table together with the previous included studies.[6–13] Authors of included studies were contacted for additional information, if necessary. Outcomes occurring relatively frequently were analyzed by calculating the pooled risk ratio (RR) with 95% confidence interval (CI) in accordance with a random-effect model. Rare outcomes with incidence <1% were combined using Peto odds ratio and a fixed-effect model.[18] All computations were performed using REVIEW MANAGER (REVMAN, Version 5.4.1 Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2020). Forest plots intended for publications were prepared using R software (Version 4.3.0, Vienna: R.

Foundation for Statistical Computing, 2023) and the forest plot package.[19,20]

To assess the risk of random errors and false-positive results and to help clarify the need for additional trials by calculating an optimal information size,[16] we performed post hoc TSA for selected outcomes in TSA viewer (Version 0.9.5.10 beta. Copenhagen: Copenhagen Trial Unit, 2017).[21]

We did not perform any subgroup analysis but conducted sensitivity analysis where we excluded one trial using old STAN technology[11] and one trial that used a different algorithm for interventions.[7] Separate analyses were prepared to explore the impact of pooling data on neonatal and perinatal deaths.

We present our overall assessment of the quality of evidence in a "summary of findings" table. The assessment includes the magnitude of effect of the STAN method vs CTG alone, and a summary of available data on the most important outcomes.[22] The quality of evidence was judged as either high, moderate, low or very low.[23]

3 | RESULTS

The new electronic searches identified 282 records; after screening of titles and abstracts, 16 records were assessed in full text, 13 were excluded and three included in the systematic review[24–26] (Figure 1 and Table 1). Reasons for exclusions and bibliography of excluded records are shown in Appendix S2. Additional and corrected data are shown in Appendix S3.

Our previous systematic review included six studies, thus data from nine randomized systematic trials were included in our updated review.

3.1 | Description of included studies

The new studies included were performed in Spain,[24] Denmark[25] and Australia,[26] with 200, 1005 and 970 women and their babies, respectively (Table 1). In all, 28 729 women and their babies were included in the updated systematic review.

Most trials used the STAN S21 or S31 monitors (Neoventa AB), whereas the Westgate trial (11) used an older device without computerized assessment for the fetal ECG (STAN 8801, Cinventa AB). The Westgate study included women from 34 weeks' gestation, and we therefore performed sensitivity analyses without that study. The decision algorithm used in the Belfort study[7] implied that the fetal heart rate status was classified into three zones (green, red, yellow), which correspond closely to the U.S. 2008 National Institute of Child and Human Development criteria.[27] If the fetal heart rate pattern is in the yellow zone, intervention is recommended if any ST event (either episodic or baseline increase) or two biphasic ST events occur. This algorithm is different from the one used in other studies. Moreover, in all studies except the Belfort study,[7] fetal blood sampling was performed in both arms at the discretion of the obstetrician in charge. Therefore, we also conducted sensitivity analysis without the Belfort study.

We assessed the overall risk of bias as low in all the included trials (Table 1, Information Appendix S4).

3.2 | The effect of STAN method vs CTG alone

The nine available trials included 28 729 women in labor but only a minority of the investigated outcomes reached statistical significance (Table 2, Appendix S5). Some of the investigated neonatal outcomes are rare, with incidences <1%, and it is difficult to gain statistical power for definite conclusions. Lack of power was not an issue for the investigated maternal outcomes, and our meta-analysis showed that STAN is associated with no difference in the rate of cesarean sections (RR 0.94; 95% CI 0.80–1.12) or assisted vaginal deliveries (RR 0.99; 95% CI 0.83–1.19) for fetal distress (Table 3).

Metabolic acidosis occurred with an incidence <1% in the group receiving CTG alone, and even lower in the STAN group (OR 0.66,

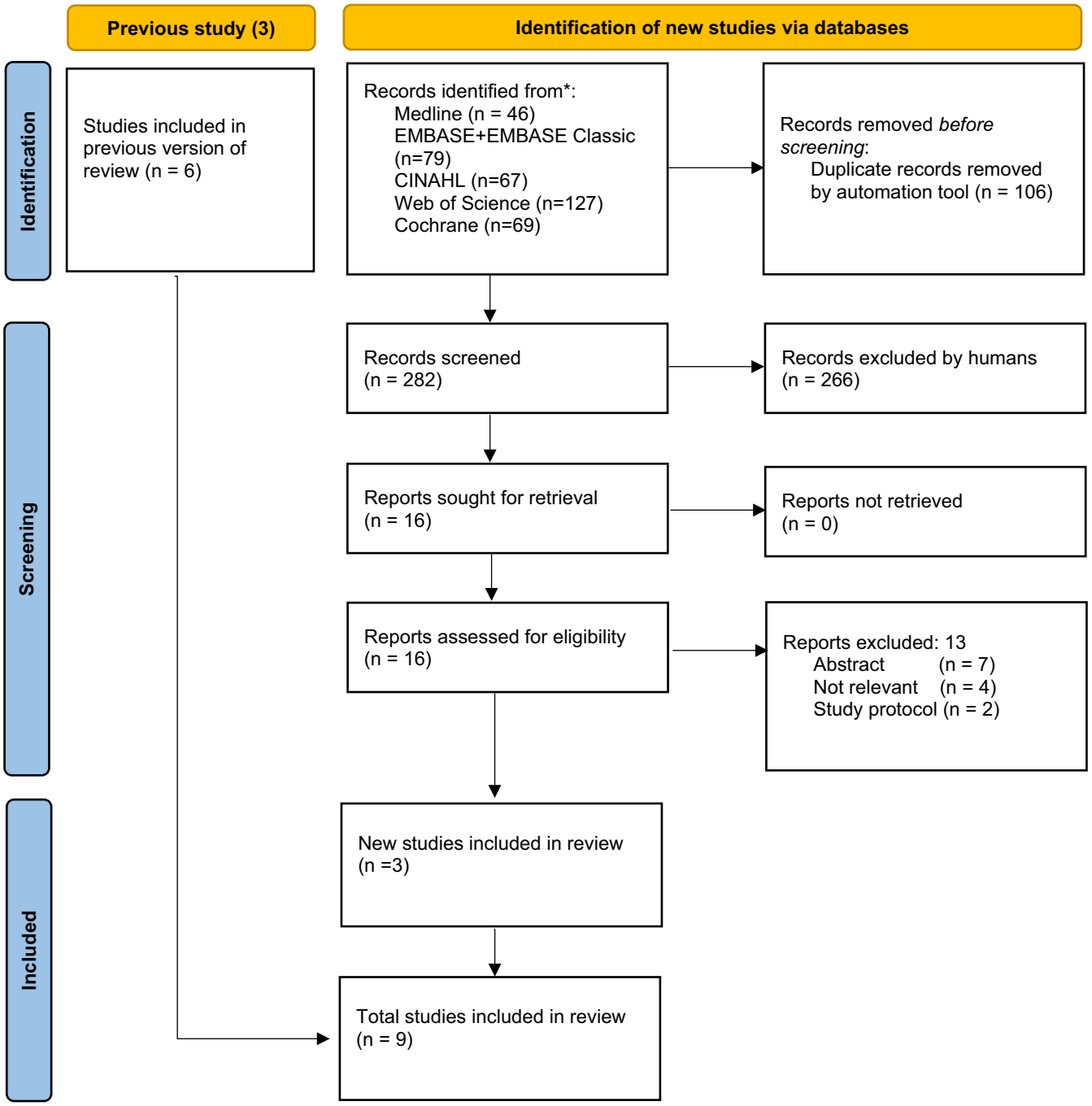

FIGURE 1 Flow diagram of the study selection process.

95% CI 0.48–0.90; Table 3). The difference corresponds to a number needed to treat to benefit of 441 (95% CI 249–1898) when the baseline risk is 0.7%. This means that one case of neonatal metabolic acidosis is avoided for every 441 women monitored with STAN instead of conventional CTG.

All included studies reported data on deaths and four reported neonatal seizures (Figure 2). Neither resulted in statistically significant differences between the STAN method vs CTG alone. Perinatal and neonatal deaths had an OR of 1.55 (95% CI 0.62–3.91) and neonatal seizures 0.86 (95% CI 0.29–2.56). The CIs were wide when expressed in relative terms, but re-expressed in absolute terms, they imply that STAN can be associated with two fewer to 14 more deaths per 10000 births, and between seven fewer and 15 more neonatal seizures per 10000 births (Table 3). Apgar scores <4 after 5 minutes seemed to occur more frequently with STAN (OR 2.86, 95% CI 1.13–7.24) but we found little or no difference with regard to the incidence of newborns with Apgar scores <4 after 1 minute (RR 1.11, 95% CI 0.61–1.99) and Apgar scores <7 after 5 minutes (RR 0.89, 95% CI 0.69–1.15; Table 3, Figure 2).

Of the investigated maternal outcomes, only fetal blood sampling reached statistical significance; this occurred less frequently in the STAN (12.5% vs 19.6%; RR 0.62, 95% CI 0.49–0.80). The

magnitude of this effect was inconsistent across the available trials (Appendix S5). Similarly, the other investigated neonatal outcomes pointed towards little or no difference between STAN vs CTG alone.

3.3 | Sensitivity analyses

The results are robust with regard to inclusion or exclusion of the Westgate[11] or Belfort[7] trials (Appendix S5). Because we pooled studies reporting perinatal and neonatal deaths, we also conducted a sensitivity analysis to explore the impact of this decision. The remaining results were consistent (Appendix S5).

3.4 | Trial sequential analyses

We decided that a relative risk reduction of 20% would represent a clinically important difference in the number of operative deliveries for fetal distress (cesarean sections, vacuum or forceps). In this case, the TSA suggested that the available data was sufficient to conclude that the two treatments are non-inferior (Appendix S6). Furthermore, as the majority of newborns with metabolic acidosis are without symptoms or elevated risk for adverse outcomes,[28,29] we held 50% relative risk reduction as the clinically important difference in the incidence of metabolic acidosis. The main analysis indicated that there was a statistically significant difference between STAN and CTG alone (Appendix S6) but the conclusion depended on the choice of statistical methods. For example, the significance was lost when we used Peto OR in combination with a random-effect model rather than in combination with a fixed-effect model. With regard to perinatal and neonatal deaths and neonatal seizures, the results were far from statistically significant, but the number of observed events was too small to allow firm conclusions about superiority or non-inferiority. For Apgar score <7 at 5 minutes, TSA confirmed there were no important differences between the groups.

3.5 | Summary of findings

The application of GRADE showed that the quality of evidence was moderate or high for the selected outcomes (Table 3).

4 | DISCUSSION

In this updated systematic review and meta-analysis of randomized controlled trials comparing ST waveform analysis against CTG alone, we found no significant difference in operative deliveries for fetal distress (either for cesarean sections or assisted vaginal deliveries) but there was a reduction in metabolic acidosis. We found no difference in neonatal and perinatal deaths, neonatal seizures or encephalopathy, transfers to NICU or Apgar score <7 at 5 minutes, or in the composite outcomes. The number of fetal scalp blood samples were significantly reduced in the STAN group compared with the CTG group. No significant differences were found in cesarean section rates or assisted vaginal deliveries for any indications.

The updated review included three new studies[24–26] with 2175 women, and thus nine randomized trials with 28 729 women and their babies were included. The updated review shows similar results as the previous version,[3] except that a previously reported significant reduction in the number of operative vaginal deliveries for any indications following STAN group disappeared.

Our review has several strengths. The findings are based on a thorough and up-to-date literature search that includes all available RCTs. All trials are associated with a low risk of bias and our findings seem robust regarding the sensitivity analyses, where we excluded two trials that prompted questions regarding external validity.[7,11] In addition, we used GRADE to judge the quality of the evidence and TSA to assess statistical power considerations for different outcomes.

RCTs are considered the gold standard for clinical trials. They are typically conducted under the supervision of dedicated experts and in ideal conditions. Therefore, the external validity to a normal clinical setting (the distinction between efficacy and effectiveness), can be questioned. The setting is almost never identical across all trials investigating the effect of an intervention, and this was also the case for the nine available STAN trials. We believe the observed variation in settings is as close as can be expected to normal variation in clinical practice, and therefore we decided to include all nine trials in our meta-analysis. However, we are aware that the appropriateness of including some of the trials has been discussed.[30–38] We therefore conducted sensitivity analyses to investigate the robustness of our results. The overall conclusions of this review are robust with regard to the inclusion or exclusion of the oldest study that used non-computerized ST analysis[11] and the study from USA that used another decision algorithm.[7]

Of the numerous outcomes reported in the included trials, we argue that the most important are perinatal and neonatal death, neonatal encephalopathy, seizures and Apgar score <4 at 5 minutes. Important long-term neurologic sequelae such as cerebral palsy or other neurodevelopmental morbidity are unfortunately not reported. Outcomes such as Apgar score <7 at 5 minutes, intubation for ventilation and transfers to NICU are less important. From a methodological perspective, we note that all relevant neonatal outcomes occur with very low incidence (for example, <0.1% for death and 0.56% for metabolic acidosis). Under such circumstances, there is a risk that the use of relative effect sizes such as odds ratios inflates the reader's perception of the magnitude of a possible effect.[39] Misconception can be avoided by presenting the relative effect sizes together with the corresponding difference in absolute terms (Table 3). The absolute risk reduction in metabolic acidosis in the STAN group compared with the CTG group is 0.23% and is probably of little clinical relevance, although the relative risk reduction is 34%.

TABLE 1 Characteristics of included studies.

Paper	Amer-Wåhlin, Sweden[6,12]	Belfort, USA[7]	Kuah, 2023[26]	Ojala, Finland[8]	Puertas, Spain[24]
Type of study	**Multicenter (3 centers)**	**Multicenter (16 centers)**	**Single center**	**Single center**	**Single center**
No. included	5049 (revised ITT-analyses)	11 108	970	1483	200
Inclusion criteria	Women in active labor ≥36 GW, with singleton fetuses in a cephalic presentation and where a clinical decision of continuous internal CTG	Women with a singleton fetus at more than 36 GW who were attempting vaginal delivery and had cervical dilation of 2 to 7 cm	Women with a singleton fetus in a cephalic position, greater than or equal to 36 GW, with naturally or artificially ruptured amniotic membranes. It was expected that labor was going to establish and progress, or they were in established labor up until active second stage of labor.	Consecutive women in active labor with term (≥36+0 GW) pregnancy, with a singleton fetus in a cephalic presentation and a decision about amniotomy	Women with late-term pregnancy (gestational age between 291 and 294 days) with a singleton pregnancy and cephalic presentation in the active stage of labor.
Exclusion criteria		Noncephalic presentation, planned CS, a need for immediate delivery, absent fetal heart-rate variability or a sinusoidal pattern, minimal fetal heart-rate variability in the 20 minutes before randomization, or other fetal or maternal conditions that would preclude a trial of labor or the placement of a scalp electrode	< 18 years old, fetal scalp electrode contraindicated, fetal cardiac abnormalities and whether the woman had participated in the study in a previous pregnancy	Contraindications for scalp electrode or admitted to the labor ward in the second phase of labor	Women with a prior cesarean section, fetal anomalies, genital bleeding, placenta previa or maternal genital infection.
Intervention	STAN S21 (Neoventa AB)	STAN S31 (Neoventa AB)	STAN S31 (Neoventa AB)	STAN S21 (Neoventa AB)	STAN S31 (Neoventa AB)
Control	Masked STAN S21	Masked STAN S31	CTG (Philips or Neoventa S31)	CTG (Hewlett-Packard 8030A)	CTG (Philips Avalon FM30)
Algorithm for interventions in STAN group	Table 1, Amer-Wåhlin	Supplementary appendices and trial protocol, referred to in Belfort	As Amer-Wåhlin	As Amer-Wåhlin	FIGO guideline 1987 and Amer-Wåhlin 2007
Proportion primiparas	62% in both arms[a]	42.6% in both arms	60.0% in CTG + STAN arm, 59.8% in CTG arm	51.0% in CTG + STAN arm, 52.4% in CTG arm	79% in CTG + STAN arm, 63% in CTG arm
Previous CS	–	–	3.9% in CTG + STAN arm, 6.2% in CTG arm	–	None
Induction of labor	17% in both arms[a]	59.2% in CTG + STAN arm, 58.6% in CTG arm	83.0% in CTG + STAN arm, 79.2% in CTG arm	20.2% in CTG + STAN arm, 17.5% in CTG arm	94% in both arms
Epidural analgesia	37% in CTG + STAN arm, 40% in CTG arm	–	84.9% in CTG + STAN arm, 84.3% in CTG arm	54.6% in CTG + STAN arm, 54.0% in CTG arm	95% in CTG + STAN arm, 91% in CTG arm
Overall risk of bias[b]	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias

Abbreviation: GW, gestational weeks.

[a]Originally, 5049 women were included and randomized to the study. Of these, 83 were excluded for technical reasons, leaving 4966 women for the analyses. In 2011 (ref) the authors published analyses according to intention-to-treat including the 83 previous excluded cases. The estimates are based on the publication from 2001.[6]

[b]See Appendix S3 for detailed risk of bias assessment.

It is common to view metabolic acidosis as a crucial outcome, but we regard it as a surrogate endpoint. The appropriate use of surrogate endpoints requires accurate knowledge and direct correlation between the surrogate and the truly important outcome. We argue that the relationship between metabolic acidosis and harder outcomes is questionable. There is a known relationship between low

Vayssiere, France[9]	Victor, Denmark[25]	Westerhuis, The Netherlands[10,13]	Westgate, UK[11]	Westgate, UK[11]
Multicenter (2 centers)	**Multicenter (2 centers)**	**Multicenter (9 centers)**	**Single center**	**Single center**
799	1005 (primary analysis) 982 (secondary analysis)	5681	2434	2434
Women in labor with a term (≥ 36 GW) singleton fetus in cephalic presentation who met the following inclusion criteria: abnormal CTG or thick meconium-stained amniotic fluid	All women ≥18 years old with a singleton fetus in cephalic presentation attempting vaginal delivery at more than 36+0 GW. Women who developed intermediate or pathological CTG abnormalities according to FIGO 1987 guidelines and with an FBS-pH sample of above 7.25 were eligible for inclusion.	Laboring women aged ≥18 years with a singleton high-risk pregnancy, a fetus in cephalic presentation, ≥36 GW, and an indication for internal electronic monitoring	All pregnancies of >34 GW with no gross fetal abnormality and with a decision to apply a scalp electrode	All pregnancies of >34 w gestation with no gross fetal abnormality and with a decision to apply a scalp electrode
Gestational age <36 GW, normal CTG during labor, maternal infection contraindicating placement of scalp electrodes (seropositive for HIV or hepatitis B or C) cardiac malformation, severe decelerations with variability reduced immediately on entry into the delivery room	Preterminal CTG, maternal HIV or hepatitis B or C, known severe fetal malformations, fetal arrhythmias or A-V blocks, and signs of chorioamnionitis.			
STAN S21 (Neoventa AB)	STAN S21 (Neoventa AB)	STAN S21 or S31 (Neoventa AB)	STAN 8801 (Cinventa AB)	STAN 8801 (Cinventa AB)
CTG (Hewlett-Packard 8030A)	Masked STAN S21	Conventional FHR monitor	CTG (Hewlett-Packard 8040A)	CTG (Hewlett-Packard 8040A)
As Amer-Wåhlin	FIGO guidelines 1987 and Neoventa Medical education material	FIGO guidelines and STAN clinical guidelines, see Appendices 1 & 2 in Westerhuis	Tables 1 and 2, Westgate	Tables 1 and 2, Westgate
72.2% in CTG+STAN arm, 71.8% in CTG arm	86.7% in CTG+STAN arm, 86.0% in CTG arm	57.2% in CTG+STAN arm, 57.0% in CTG arm	–	–
6.3% in CTG+STAN arm, 6.0% in CTG arm	5.6% in CTG+STAN arm, 4.5% in CTG arm	12.2% in CTG+STAN arm, 13.1% in CTG arm	–	–
8.5% in CTG+STAN arm, 8.8% in CTG arm	34.1% in CTG+STAN arm, 35.9% in CTG arm	40.9% in CTG+STAN arm, 41.8% in CTG arm	–	–
91.2% in CTG+STAN arm, 90.3% in CTG arm	88.0% in CTG+STAN arm, 83.4% in CTG arm	41.7% in CTG+STAN arm, 42.6% in CTG arm	–	–
Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias

cord artery pH values and serious outcome, but the threshold remains unknown.[40,41] Few neonates with severe acidemia appear to have sequelae, particular those who are healthy at birth, and most neonates with adverse outcomes, also those with seizures, are not born acidemic.[28,42] Recent studies also report that umbilical artery pH and base excess are poor predictors of short-term outcomes

TABLE 2 Outcome events and meta-analyses.

Outcome	No. of studies	Events, *n/N*	Effect measure[a]	Effect size (95% CI)	I^2 (%)
Metabolic acidosis	9	160/28342	Peto OR	0.66 (0.48–0.90)	28
Total operative deliveries[b] for fetal distress	9	3157/28618	RR	0.96 (0.82–1.11)	80
Perinatal and neonatal deaths	9	18/28918	Peto OR	1.55 (0.62–3.91)	0
Neonatal seizures	4	13/14310	Peto OR	0.86 (0.29–2.56)	26
Apgar score <4 at 1 minute	2	42/1204	RR	1.11 (0.61–1.99)	0
Apgar score <4 at 5 minutes	2	23/12020	Peto OR	2.86 (1.13–7.24)	–
Apgar score <7 at 5 minutes	8	238/17473	RR	0.89 (0.69–1.15)	0
Neonatal encephalopathy	5	25/24144	Peto OR	0.66 (0.30–1.46)	14
Neonatal intubation	4	97/13711	Peto OR	1.44 (0.96–2.14)	0
Fetal blood sampling	8	2814/17510	RR	0.62 (0.49–0.80)	91
Admittance to NICU	9	1658/28582	RR	1.00 (0.91–1.10)	0
Umbilical cord pH <7.05	4	216/10336	RR	1.05 (0.63–1.76)	66
Composite endpoint[c]	2	102/12075	Peto OR	1.33 (0.90–1.96)	0
Composite endpoint[d]	1	83/1005	Peto OR	0.99 (0.63–1.56)	–
Total operative deliveries[b] for all indications	9	7643/28618	RR	0.99 (0.94–1.04)	34
Cesarean deliveries for fetal distress	9	1389/28618	RR	0.94 (0.80–1.12)	54
Cesarean delivery for all indications	9	4206/28618	RR	1.00 (0.95–1.06)	0
Operative vaginal delivery for fetal distress	9	1798/28618	RR	0.99 (0.83–1.19)	71
Operative vaginal deliveries for all indications	9	3440/28618	RR	0.99 (0.89–1.09)	56

[a]Peto odds ratio (OR), fixed effect model, for outcomes with incidence <1%, otherwise risk ratio (RR), random effect model.

[b]Total operative deliveries = cesarean sections + assisted vaginal deliveries.

[c]Composite of intrapartum death, neonatal death, Apgar <4 at 5 minutes, neonatal seizures, metabolic acidosis, intubation at birth, or neonatal encephalopathy.

[d]Composite endpoint of 1-minute Apgar score <4 or 5-minute Apgar score <7 or metabolic acidosis or admission to NICU >24 hours.

TABLE 3 Summary of findings for selected outcomes.

Outcomes	Anticipated absolute effects[a] (95% CI)		Relative effect (95% CI)	No. of participants (studies)	Quality of the evidence (Grade)
	Risk with CTG alone	Risk with STAN			
Metabolic acidosis cord pH <7.05 + $BD_{(ecf)}$ >12 mmol/L	68 per 10000	45 per 10000 (33–61)	OR 0.66 (0.48–0.90)	28342 (9 RCTs)	⨁⨁⨁◯ Moderate[b,c]
Cesarean section for fetal distress	493 per 10000	493 per 10000 (469–523)	RR 1.00 (0.95–1.06)	28618 (9 RCTs)	⨁⨁⨁⨁ High
Operative vaginal delivery for fetal distress	619 per 10000	612 per 10000 (513–736)	RR 0.99 (0.83–1.19)	28648 (9 RCTs)	⨁⨁⨁⨁ High
Neonatal and perinatal death	5 per 10000	8 per 10000 (3–19)	OR 1.55 (0.62–3.91)	28618 (9 RCTs)	⨁⨁⨁◯ Moderate[d,e]
Neonatal seizures	10 per 10000	8 per 10000 (3–25)	OR 0.86 (0.29–2.56)	14310 (4 RCTs)	⨁⨁⨁◯ Moderate[d,e]
Apgar score <7 at 5 minutes	145 per 10000	129 per 10000 (100–166)	RR 0.89 (0.69–1.15)	17473 (8 RCTs)	⨁⨁⨁⨁ High

[a]The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[b]Inconsistency in effect estimates.

[c]A surrogate endpoint with questionable clinical importance. Choose not to downgrade.

[d]Wide confidence intervals – imprecise data.

[e]Optimal information size not achieved.

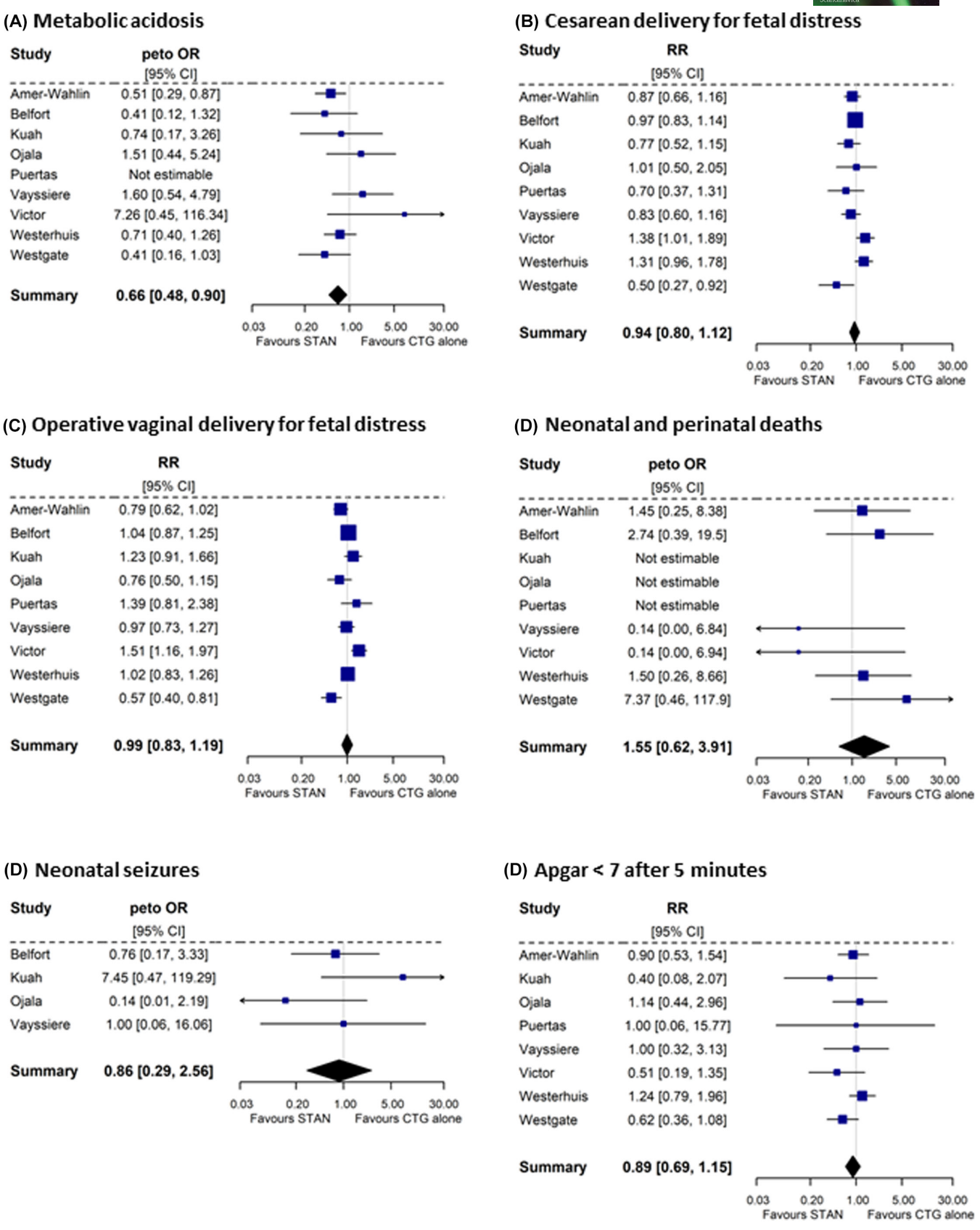

FIGURE 2 Forest plot analyses for selected outcomes.

as low 5-minute Apgar score and long-term neurodevelopmental morbidity.[29,42,43]

The causes of severe long-term neurologic sequelae are probably more complex than previously believed and not simply due to hypoxia with metabolic acidosis.[44] Thus, it seems obvious that metabolic acidosis is a surrogate endpoint and should be interpreted with caution. We found a statistically significant difference in favor of STAN when comparing the incidence of metabolic acidosis, without observing similar effects in other important outcomes.

In addition to conventional meta-analysis, we performed TSA on selected outcomes to explore the possible impact of random errors and false-positive results on the conclusions of our meta-analysis. TSA also allow power analysis to clarify the need for additional trials.[16] These analyses showed that the statistical power is too low to draw firm conclusions about superiority or non-inferiority of either STAN or CTG alone for neonatal seizures or deaths. On the other hand, TSA showed adequate statistical power to conclude that the STAN method is probably not associated with important reductions in Apgar <7 at 5 minutes or with operative deliveries for all indications (cesarean sections, vacuum or forceps).

We found that metabolic acidosis was associated with a statistically significant improvement in favor of STAN. REVMAN does not enable the use of a random-effect model in combination with Peto OR effect sizes, therefore the main analysis was based on a fixed-effect model. Interestingly, the TSA analysis showed that the significant finding for metabolic acidosis disappeared in meta-analysis based on random-effect models, a result that underpins the need for caution in interpreting the statistically significant finding for metabolic acidosis.

A recent Norwegian population study investigated the impact of the introduction of the STAN technology on changes in the occurrence of fetal and neonatal deaths, Apgar scores of <7 at 5 minutes, intrapartum cesarean sections and assisted vaginal deliveries while controlling for time- and hospital-specific trends and maternal risk factors.[45] The analyses found that the introduction of STAN into clinical practice had no impact on fetal or neonatal deaths, either on the rates of cesarean section or assisted vaginal deliveries. However, it was associated with a small, but statistically significant increase in the occurrence of babies with Apgar score <7 at 5 minutes.[45]

The reduction in fetal blood sampling was expected, as it is one main reason for introducing the STAN technology, although fetal blood sampling was optional in most of the RCTs.

In a recent commentary, Chandraharan stated that the problem with STAN is the manufacturer's guidelines for interpreting CTG grouped into the classes "normal", "intermediary" and "abnormal".[46] He argued that if a physiological interpretation of CTG was used instead of the current tool based on pattern recognition, STAN's true potential to improve clinical outcomes might be observed.[46] The theory that implementing a new guideline of physiological interpretation of CTG compared with current guidelines based on pattern recognition will improve clinical outcomes, remains to be tested through clinical trials.

Today, the STAN method is in widespread use in Denmark and Norway, but not in Sweden or Iceland. In Finland, one obstetric unit uses STAN. It is used in 20% of the obstetric units in UK, none in Ireland, and in some units in the Netherlands and Belgium and some other European countries. STAN is used in one hospital in Australia, and not used in the USA.

5 | CONCLUSION

Our updated systematic review and meta-analysis of nine randomized controlled trials comparing ST waveform analysis against CTG alone, including 28 729 women and their babies, showed no reduction in important clinical outcomes such as severe neonatal morbidity, mortality rates or operative delivery rates. The significant but modest absolute reduction of metabolic acidosis of 0,23% should be interpreted with caution. To our best knowledge, no new randomized clinical trial is planned and it is time to conclude that STAN carries no important clinical benefits compared with CTG alone.

AUTHOR CONTRIBUTIONS

EB screened titles and abstracts, assessed articles in full text, assessed risk of bias, extracted data, graded the results and wrote the first draft. KGB performed the analyses, wrote the Method section and graded the results. ER performed the literature searches and described the searches in the paper. LMR assessed risk of bias and graded the results. PØ screened titles and abstracts, assessed articles in full text and extracted data. All authors contributed to revision of the paper.

CONFLICT OF INTEREST STATEMENT

None declared.

ORCID

Ellen Blix https://orcid.org/0000-0001-7971-4580

REFERENCES

1. Alfirevic Z, Devane D, Gyte GM, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev*. 2017;2017:CD006066.
2. Rosen KG. Fetal electrocardiogram waveform analysis in labour. *Curr Opin Obstet Gynecol*. 2005;17:147-150.
3. Blix E, Brurberg KG, Reierth E, Reinar LM, Øian 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.
4. Saccone G, Schuit E, Amer-Wåhlin I, Xodo S, Berghella V. Electrocardiogram ST analysis during labor: a systematic review and meta-analysis of randomized controlled trials. *Obstet Gynecol*. 2016;127:127-135.
5. Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. *Cochrane Database Syst Rev*. 2015;2015:CD000116.

6. Amer-Wåhlin I, Hellsten C, Norén H, et al. Cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram for intrapartum fetal monitoring: a Swedish randomised controlled trial. *Lancet*. 2001;358:534-538.
7. Belfort MA, Saade GR, Thom E, et al. A randomized trial of intrapartum fetal ECG ST-segment analysis. *New Engl J Med*. 2015;373:632-641.
8. Ojala K, Vääräsmäki M, Mäkikallio K, Valkama M, Tekay A. A comparison of intrapartum automated fetal electrocardiography and conventional cardiotocography—a randomised controlled study. *BJOG*. 2006;113:419-423.
9. Vayssiere 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(299):e1-e6.
10. Westerhuis ME, Visser GH, Moons KG, et al. Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. *Obstet Gynecol*. 2010;115:1173-1180.
11. Westgate J, Harris M, Curnow JS, Greene KR. Randomised trial of cardiotocography alone or with ST waveform analysis for intrapartum monitoring. *Lancet*. 1992;340:194-198.
12. Amer-Wåhlin I, Kjellmer I, Marsál K, Olofsson P, Rosen 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;90:990-996.
13. Westerhuis ME, Visser GH, Moons KG, Zuithoff N, Mol BW, Kwee A. Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. *Obstet Gynecol*. 2011;117:406-407.
14. Al Wattar BH, Honess E, Bunnewell S, et al. Effectiveness of intrapartum fetal surveillance to improve maternal and neonatal outcomes: a systematic review and network meta-analysis. *CMAJ*. 2021;193:E468-E477.
15. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924-926.
16. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol*. 2008;61:64-75.
17. Higgins JPT, Savoviv J, Page MJ, Elbers R, Sterne JAC. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.4*. The Cochrane Collaboration; 2023 Available online at Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training.
18. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3*. The Cochrane Collaboration; 2022 Available online at: www.training.cochrane.org/handbook
19. Team RC. *Team R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2014 Available online at: http://www.R-project.org
20. Gordon M, Lumley T. Forestplot: Advanced Forest Plot Using "grid" Graphics. R package version 3.1.3. Available online at: http://CRAN.R-project.org/package=forestplot
21. Thorlund KEJ, Wetterslev J, Brok J, Imberger G, Gluud C. *User Manual for Trial Sequential Analysis (TSA)*. 2nd ed. Copenhagen Trial Unit; 2017.
22. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64:383-394.
23. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401-406.
24. Puertas A, Góngora J, Valverde M, Revelles L, Manzanares S, Carrillo MP. Cardiotocography alone vs cardiotocography with ST segment analysis for intrapartum fetal monitoring in women with late-term pregnancy. A randomized controlled trial. *Eur J Obstet Gynecol R B*. 2019;234:213-217.
25. Victor SF, Bach DBB, Hvelplund AC, et al. Cardiotocography combined with ST analysis versus cardiotocography combined with fetal blood sampling in deliveries with abnormal CTG: a randomized trial. *Arch Gynecol Obstet*. 2023;307:1771-1780.
26. Kuah S, Simpson B, Salter A, et al. Comparing the effect of CTG+STan with CTG alone on emergency cesarean section rate: STan Australian randomized controlled trial (START). *UOG*. 2023;62:462-470. doi:10.1002/uog.26279
27. Macones GA, Hankins GD, Spong CY, Hauth J, Moore T. The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. *Obstet Gynecol*. 2008;112:661-666.
28. Hafström 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-e1507.
29. Leinonen E, Gissler M, Haataja L, et al. Umbilical artery pH and base excess at birth are poor predictors of neurodevelopmental morbidity in early childhood. *Acta Paediatr*. 2019;108:1801-1810.
30. Potti S, Berghella V. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: a meta-analysis of randomized trials. *Am J Perinatol*. 2012;29:657-664.
31. Becker JH, Bax L, Amer-Wåhlin I, et al. ST analysis of the fetal electrocardiogram in intrapartum fetal monitoring: a meta-analysis. *Obstet Gynecol*. 2012;119:145-154.
32. 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.
33. Schuit E, Amer-Wåhlin 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(187):e1-e13.
34. 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;93:571-586.
35. 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 I: the randomized controlled trials. *Acta Obstet Gynecol Scand*. 2014;93:556-568.
36. Øian P, Blix E. Scarce scientific evidence for the use of cardiotocography plus fetal ECG ST interval analysis (STAN). *Acta Obstet Gynecol Scand*. 2014;93:570.
37. Blix E, Øian P. Deviations from STAN guidelines are frequent but results cannot be excluded when the effectiveness of the method should be evaluated. *Acta Obstet Gynecol Scand*. 2014;93:589.
38. Steer PJ, Hvidman LE. Scientific and clinical evidence for the use of fetal ECG ST segment analysis (STAN). *Acta Obstet Gynecol Scand*. 2014;93:533-538.
39. Marshall KG. Prevention. How much harm? How much benefit? 1. Influence of reporting methods on perception of benefits. *CMAJ*. 1996;154:1493-1499.
40. 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.
41. Myrhaug HT, Kaasen A, Pay ASD, et al. Umbilical cord blood acid-base analysis at birth and long-term neurodevelopmental outcomes in children: a systematic review and meta-analysis. *BJOG*. 2023;130:1156-1166.

42. Yeh P, Emary K, Impey L. The relationship between umbilical cord arterial pH and serious adverse neonatal outcome: analysis of 51,519 consecutive validated samples. *BJOG*. 2012;119:824-831.
43. Johnson GJ, Salmanian B, Denning SG, Belfort MA, Sundgren NC, Clark SL. Relationship between umbilical cord gas values and neonatal outcomes: implications for electronic fetal heart rate monitoring. *Obstet Gynecol*. 2021;138:366-373.
44. Leviton A. Why the term neonatal encephalopathy should be preferred over neonatal hypoxic-ischemic encephalopathy. *Am J Obstet Gynecol*. 2013;208:176-180.
45. Blix E, Eskild A, Skau I, Grytten J. The impact of the introduction of intrapartum fetal ECG ST segment analysis. A population study. *Acta Obstet Gynecol Scand*. 2022;101:809-818.
46. Chandraharan E. Fetal electrocardiograph (ST-Analyser or STAN): is it time for the requiem? *J Clin Med Surg*. 2023;3:1111.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Blix E, Brurberg KG, Reierth E, Reinar LM, Øian P. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: An updated systematic review and meta-analysis of randomized trials. *Acta Obstet Gynecol Scand*. 2023;00:1-12. doi:10.1111/aogs.14752