Bigger is not always better…….

The validity of the US randomized trial of STAN for Norway

Summary
The results of the US randomized trial cannot be transferred to Norwegian/European standards:
- Selection of a low-risk population (see paragraph 1)
- Other guidelines for both CTG and ST Analysis (see paragraph 2, 3 and 4)
- Guidelines with a longer intervention time at fetal distress (see paragraph 5)
- Slow learning curve / recruitment (see paragraph 6)
- Lack of strength (see paragraph 7)

History
Five randomized trials have so far been conducted in Europe with a total of more than 15,000 women who were randomized to either monitoring with CTG alone or CTG and ST Analysis (STAN) [1-5]. Fetal blood sampling was accepted and used in both arms in all the trials. Use of STAN reduced need for fetal blood sampling, as well as a decrease in vaginal operative deliveries and admission to neonatal intensive care unit.
The largest randomized studies have also shown a reduction in the prevalence of metabolic acidosis when using STAN [6].
A number of clinical observational studies from different countries, including Norway, have demonstrated a reduction in the incidence of metabolic acidosis after introducing STAN [7-9].
STAN has been in use in Norway for over 20 years and most high risk pregnancies in Norway have been monitored with STAN during the last 10 years.

It was in the involvement and the introduction of STAN that a structured training and certification of CTG was performed at all delivery wards in Norway. STAN clinical guidelines defined for the first time an intervention time at fetal distress.
Indications for use of STAN in Norway are:
- Gestational age ≥36 weeks
- Active labour after membranes are ruptured or artificial rupture of membranes
- Indication for continuous fetal monitoring due to high risk
- Initiation of STAN as early as possible and no later than the end of the first stage of labour
- Defined intervention time of birth in 1st and 2nd stage of labour

STAN technology was approved in 2005 for clinical use in the US by the FDA.

Obstetric practice in US
Over 30% of all births are delivered by caesarean section partly due to a huge medicolegal pressure. The majority of births in the US are monitored with CTG [10]. In Norway the corresponding number is estimated to be 50%.
In the US approximately 20% of nulliparous are delivered by Emergency cesarean section. Only 3.5% of the deliveries end with vaginal operative delivery.
Also fetal monitoring in the US differs significantly from that in Europe, including the use of another CTG classification. Fetal blood sampling in labour is not used in the US.

About the US randomized trial CTG vs. CTG + ST Analysis [11]
Recruitment of the trial took place from November 2010 to April 2014. Over 11,000 women were randomized.
The main inclusion criteria were:
- Cephalic presentation

- Gestational age ≥ 36 weeks

- Cervical dilatation between 2 and 7 cm

The main outcome was a "composite outcome", i.e. the occurrence of one or more of the following:
- Intrapartum/neonatal death

- Apgar <3 at 5 minutes

- Metabolic acidosis

- Need for intubation

- Neonatal seizures or encephalopathy

No difference was found in either primary or secondary outcome (e.g. occurrence of cesarean section or vaginal operative delivery) in relation to the surveillance method.

What limits the external validity?
1. _The population of the US trial is dominated by low-risk deliveries_ since the only inclusion requirement except gestational age or fetal presentation was a cervical dilatation between 2 and 7 cm. Women who have previously undergone cesarean section were excluded from the trial. No information regarding what sort of risk pregnancy the women had is available. Two of three women had given birth vaginally before (Robson group 3 and 4a), the average age was 27 years and women had an average pre-pregnant BMI of 27. All this indicates a low risk population, however there was still an induction rate of 58% for all deliveries.
The potential advantage of fetal monitoring (STAN or another method), is therefore most beneficial in a population of high risk for intervention and adverse neonatal outcomes. Opposite to the US RCT, the majority of the European RCT:s recruited high risk pregnancies to STAN. It is also in this population STAN is recommended for use in Norway.

2. The US study used _a different interpretation classification for CTG._ The biggest difference is that intermediary (yellow, Fig. 2) and pathological CTG (orange, Fig.2) were merged into one category - "yellow zone" (Fig.1)

FHR Classification	Baseline Heart Rate	Variability	Decelerations
Green Zone	• 110-160 bpm	• Moderate variability (6-25 bpm) • Accelerations present or absent	• Early decelerations: present or absent • Variable decelerations: absent or if present variable decelerations with a duration of <60 sec and depth <60 beats
Yellow Zone	• Bradycardia <110 bpm • Tachycardia >160 bpm • >150 bpm with minimal variability	• Minimal variability (>undetectable and ≤5 bpm) for >40 min • Marked variability (>25 bpm) for >40 min	• Variable decelerations with a duration of ≥60 sec or depth ≥60 beats • Recurrent late decelerations • Prolonged deceleration for >2 min regardless of variability
Red Zone	• Absent variability (undetectable) regardless of other FHR patterns • Sinusoidal pattern		

Fig. 1: CTG classification used in the US trial

Classification of CTG
Composed by the Danish and Norwegian reference group according to FIGO guidelines, December 2007

	Baseline heart frequency	Variability Reactivity	Decelerations
Normal CTG	• 110–150 bpm	• Accelerations • 5–25 bpm	• Early uniform decelerations • Uncomplicated variable decelerations (loss of <60 beats)
Intermediary CTG	• 100–110 bpm • 150–170 bpm • Short bradycardia episode <100 bpm for >3 min <80 bpm for >2 min	• >25 bpm (saltatory pattern) • <5 bpm >40 min	• Uncomplicated variable decelerations (loss of >60 beats)
	• A combination of 2 or several intermediary observations will result in an abnormal CTG		
Abnormal CTG	• >170 bpm • Persistent bradycardia <100 bpm for >10 min <80 bmp for >3 min (without an increasing tendency)	• <5 bpm for >60 min • Sinusoidal pattern	• Complicated variable decelerations with a duration of >60 sec • Repeated late uniform decelerations
Preterminal CTG	• Total lack of variability (<2 bpm) and reactivity with or without decelerations or bradycardia		

Fig. 2: CTG classification used in Norway and Europe

3. _Different STAN guidelines were used_ in the US trial. Intervention limits used in Norwegian clinical guidelines for pathological CTG were applied to all types of CTG changes, which falls into the "yellow zone". This means that ST Events in our interpretation is without importance because CTG was intermediary (continued observation) in the US trial it demanded action/intervention. Example: A Baseline T/QRS-rise of 0.06 with a tachycardia of 165 beats/min the US guidelines indicate intervention, while the Norwegian clinical guidelines allow further observation.
The occurrence of hypoxia and acidosis with an intermediary CTG is basically very low,[12] but if you use an even lower threshold for intervention (as was done in the US study) it must result in a large number of unnecessary interventions.

4. If the midwife suspected severe fetal hypoxia either by clinical changes and/or CTG changes, the study protocol in _the US trial allowed intervention in the STAN arm with CTG changes alone,_ completely independent of the reading by the ST Analysis.
A general opening of this possibility is problematic because participants in the STAN arm in practice risk to be monitored by a method they are not randomized to.

5. It allowed a significantly _longer time interval from the indication for intervention to delivery._ Unlike FDA approved guidelines should the delivery in the US trial be completed within 60 minutes in the first stage of labour and within 30 minutes in active second stage of labour. Similar interval in Norway is respectively 20 min for the the first stage and "immediate (as quick as possible) delivery" in active second stage of labour. It is shown that an intervention interval longer than 20 min increases the risk of adverse neonatal outcomes. [13]

6. Strength calculation of the US STAN randomized trial was based on a previous study of pulse oximetry [14] where only nulliparous were recruited. In this trial occurred the primary combined outcome in 1.75% and 26% were delivered by Caesarean section. In the US trial, the primary outcome occurred in 0.9% and only 17% were delivered by caesarean section. This means that the study did _not have enough strength_ in the incidence of either primary outcome or incidence of cesarean section.

7. Each participating department recruited an average of 5 women per month to the STAN arm in the US RCT (5532 women / 26 units / 41 months) [11]. Several weeks went by between each time the individual midwife had to deal with a delivery with the STAN method. Such a _low recruitment rate must have caused a very slow learning curve._ The learning curve has in previous RCT with STAN shown to have effect on the outcome [15]. In comparison with the Swedish RCT [2] was the recruitment speed 45 women/month/department.

Other comments:
This was the first RCT with a combined outcome variable instead of metabolic acidosis alone. Fetal monitoring is used to detect fetal hypoxia with risk of developing acidosis. Collective outcomes beyond metabolic acidosis is low Apgar scores, need for intubation and neonatal seizures that will include a variety of conditions, which cannot be diagnosed with CTG and/or STAN.
In this context, it is interesting that the frequency of metabolic acidosis was 2.5 times greater in the CTG arm (3 vs. 8, p = 0.13), while there were several newborn with larger malformations in the STAN arm (38 vs. 23, p = 0.05) [11].
In the US RCT with a low-risk population was 5.2% (30% of all operative deliveries) in the STAN arm delivered by cesarean section for indication fetal distress [11]. By comparison, the corresponding figure 3.2% (22.5% of all instrumental deliveries) in the Dutch RCT, as opposed to the US trial had high-risk patients been recruited only [5]. This illustrates an extremely low intervention threshold in the US trial.

What does the US Randomized Trial signify for recommendations of fetal monitoring?
It clearly shows that the STAN method does not work if one substantially modifies important clinical guidelines on which the method is based on:
- Graded intervention limits in relation to the severity of CTG changes
- Rapid intervention (20 min) where there are indications of fetal hypoxia
The study was conducted in an obstetric environment and in a population that is significantly different from the conditions in Norway.
Almost all large and medium sized delivery wards in Norway have gradually gained enormous clinical experience in the use of the STAN method and achieved good results.

Therefore there is no reason to change practices on the basis of a study which is not in any way valid for the conditions in Norway.

Norwegian reference group for fetal monitoring:
Branka Yli, Thomas Hahn, Jørg Kessler, Hilde Kristin Lie, Marit Martinussen

References:

1. Westgate, J., et al., Plymouth randomized trial of cardiotocogram only versus ST waveform plus cardiotocogram for intrapartum monitoring in 2400 cases. American Journal of Obstetrics and Gynecology, 1993. 169(5): p. 1151-60.

2. Amer-Wåhlin, I., et al., 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 Obstetricia et Gynecologica Scandinavica, 2011. 90(9): p. 990-6.

3. Ojala, K., et al., A comparison of intrapartum automated fetal electrocardiography and conventional cardiotocography - a randomised controlled study. BJOG-an International Journal of Obstetrics and Gynaecology, 2006. 113(4): p. 419-423.

4. Vayssière, C., et al., A French randomized controlled trial of ST-segment analysis in a population with abnormal cardiotocograms during labor. American Journal of Obstetrics and Gynecology, 2007. 197(3): p. 299.e1-6.

5. Westerhuis, M.E., et al., Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. Obstetrics and Gynecology, 2010. 115(6): p. 1173-80.

6. Olofsson, P., et al., A critical appraisal of the evidence for using cardiotocography plus ECG ST interval analysis for fetal surveillance in labor. Part II: the metaanalyses. Acta Obstet Gynecol Scand, 2014. 93(6): p. 571-86.

7. Doret, M., et al., Use of peripartum ST analysis of fetal electrocardiogram without blood sampling: a large prospective cohort study. European Journal of Obstetrics,Gynecology, and Reproductive Biology, 2011. 156(1): p. 35-40.

8. Norén, H. and A. Carlsson, Reduced prevalence of metabolic acidosis at birth: an analysis of established STAN usage in the total population of deliveries in a Swedish district hospital. American Journal of Obstetrics and Gynecology, 2010. 202: p.546.e1-7.

9. Kessler, J., D. Moster, and S. Albrechtsen, Intrapartum monitoring of high-risk deliveries with ST analysis of the fetal electrocardiogram: an observational study of 6010 deliveries. Acta Obstet Gyn Scan, 2013. 92(1): p. 57-84.

10. Chen, H.Y., et al., Electronic fetal heart rate monitoring and its relationship to neonatal and infant mortality in the United States. Am J Obstet Gynecol, 2011.204(6): p. 491 e1-10.

11. Belfort, M.A., et al., A Randomized Trial of Intrapartum Fetal ECG ST-Segment Analysis. N Engl J Med, 2015. 373(7): p. 632-41.

12. Melin, M., et al., Changes in the ST-interval segment of the fetal electrocardiogram in relation to acid-base status at birth. BJOG, 2008. 115(13): p. 1669-75.

13. Kessler, J., D. Moster, and S. Albrechtsen, Delay in intervention increases neonatal morbidity in births monitored with cardiotocography and ST-waveform analysis. Acta Obstetricia et Gynecologica Scandinavica, 2013.

14. Bloom, S.L., et al., Fetal pulse oximetry and cesarean delivery. N Engl J Med, 2006. 355(21): p. 2195-202.

15. Amer-Wahlin, I., et al., Implementation of new medical techniques: experience from the Swedish randomized controlled trial on fetal ECG during labor. J Matern Fetal Neonatal Med, 2005. 18(2): p. 93-100.