

FDA Draft Panel Presentation (6-17-05)

Neovanta Medical AB STAN S31 Fetal Heart Monitoring System (P020001)

Meeting Date June 23, 2005



1

STAN S31 Fetal Heart Monitor

Kathryn S. Daws-Kopp

Electrical Engineer / Lead Reviewer
Obstetrics and Gynecological Devices Branch
Division of Reproductive, Abdominal, and
Radiological Devices

CDRH-ODE



2

STAN S31 Monitor

A Specialized Perinatal Monitor
Providing Fetal ECG Analysis



3

Indication for Use

- ✦ For pregnancies at term (>36 weeks),
- ✦ To improve assessment of fetal condition during labor, and
- ✦ As adjunct to standard fetal heart monitoring

Use is indicated when there is planned vaginal delivery and:

- ✦ need for close fetal surveillance during labor, or
- ✦ maternal disorders and/or utero-placental dysfunction with potential adverse influence on fetal oxygen and nutritional supply, or
- ✦ deviation from the normal course of labor, ex: induction/augmentation



4

Review Team

- ✦ Kathy Daws-Kopp → *Lead, Electrical Safety, EMC*
- ✦ Julia Carey-Corrado, M.D. → *Clinical*
- ✦ Gene Pennello, Ph.D. → *Statistical*
- ✦ Danica Marinac-Dabic, M.D., Ph.D. → *Epidemiology*
- ✦ Baoguang Wang, M.D., Dr. P.H. → *Epidemiology*
- ✦ Sandy Weininger, Ph.D. → *Software & Design*
- ✦ Linda Godfrey → *Bioresearch Monitoring*
- ✦ Sharon Murrain-Ellerbe → *Manufacturing*



5

History of FDA Review

- ✦ January 2002—Original PMA
- ✦ April 2002—Orig PMA to Panel
- ✦ June 2003—Closed Session
- ✦ February 2005—Amendment



6

Objectives

- ◆ Device design
 - ◆ Components
 - ◆ Mechanism of Action
- ◆ Pre-clinical review focus
- ◆ Ongoing issues



7

Device Components

- Base Unit
- Monitor
- Software
- Sensors



8

Device Components-Sensors

Off-the-Shelf Components

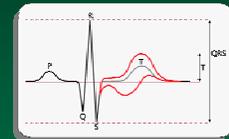
- Spiral Electrode (fetal heart rate + FECG)
- Doppler FHR (external heart rate)
- IUP (uterine activity)
- Toco (uterine activity)
- Event Marker



9

Mechanism of Action

- ◆ Responses to hypoxia show in ECG waveform
- ◆ FECG Analysis
 - ◆ T/QRS ratio (baseline or episodic)
 - ◆ ST segment (biphasic 1, 2, or 3)
 - ◆ 20 minute period - no event posting



10

Things we look at

- ◆ Software/Hardware — design / testing
- ◆ Bioresearch Monitoring — study execution
- ◆ Manufacturing — comp. w / design controls
- ◆ Clinical
- ◆ Statistical



11

Changes from STAN S21 to S31

- ◆ Human factor enhancements
- ◆ Minor hardware enhancements
- ◆ Support for connectivity options (USB and Ethernet)
- ◆ External monitoring (ultrasound FHR)



12

Changes from STAN S21 to S31

- ◆ Data acquisition & signal processing very similar
- ◆ Updated design/construction-- demonstrated to specs
- ◆ Bench testing (including simulation data)



13

Ongoing Issues

- ✓ *Bioresearch Monitoring*
- ✓ *Manufacturing*



14

Julia Carey-Corrado, M.D.

Clinical Review Issues

Gene Pennello, Ph.D.

Statistical Issues



15

Update on FDA Clinical Review of STAN S31 Fetal Heart Monitor

June 23, 2005

Julia Carey-Corrado, M.D.
CDRH/ODE/OGDB



16

Objectives

- Discuss Swedish RCT as pivotal clinical trial of STAN
- Review FDA analysis of April 2002 Panel recommendation
- Present FDA Analysis of STAN US Education and Clinical Use Study



17

Overview of Swedish RCT

- Amer-Wahlin I, Hellsten C, Noren H et al. Lancet 2001; 358:534-38
- Prospective, randomized, multi-center, controlled
- December 1998-June 2000
- 4966 labors at 3 University Hospitals in Sweden
- STAN (ST + FHR) vs control (FHR-only)



18

Swedish RCT - Study Endpoints

- Umbilical cord acid-base status
- Change in frequency of operative delivery
- Neonatal morbidity as identified by Apgar scores at 5 minutes, NICU admissions, neurologic signs and death



19

Swedish RCT Results (Intent-to-Treat)

	FHR-only	FHR + ST	P - value
Metabolic Acidosis	1.5% (31/2079)	0.7% (15/2159)	0.02
ODFD	9.3% (227/2447)	7.7% (193/2519)	0.047
CSFD	4.0% (97/2447)	3.5% (87/2519)	0.38



20

Swedish RCT - Results (minus inadequate recording cases)

	FHR-only	FHR + ST	P - value
Metabolic Acidosis	1.44% (27/1871)	0.57% (11/1926)	0.01
ODFD	7.99% (173/2164)	5.92% (132/2228)	0.009
CSFD	2.91% (63/2164)	1.93% (43/2228)	0.04



21

Swedish RCT - Neonatal Morbidity

	FHR-only N=2447	FHR + ST N=2519	P - value
5 min Apgar <7	28	26	0.81
Admission to NICU	181	169	0.21
Moderate to severe encephalopathy	8	1	0.02
Perinatal death	1	2	



22

Swedish RCT Selected FDA Review Issues

- Planned interim analysis after 1600 subjects treated
 - Study underpowered (to detect difference in metabolic acidosis)
 - Protocol deviations
- Pre-existing hypoxia
- Lack of automatic signal
 - Poor signal quality
 - Inadequate recording time



23

Summary of 2002 Panel Discussion

- Pros
 - Potential value of STAN in improving neonatal neurological outcomes
 - Safety and effectiveness demonstrated in Swedish RCT



24

Summary of 2002 Panel Discussion

- Cons
 - Lack of US clinical data
 - Classification of FHR patterns unfamiliar to US clinicians
 - Clinical Guidelines employed FIGO terminology and conventions
 - Differences in obstetrical practice between Europe and US



25

2002 Panel Recommendation

- By a 6-5 vote, Panel recommended non-approval because PMA did not demonstrate that STAN technology could be successfully transferred into obstetrical practice in the US.



26

FDA Decision June 2002

- PMA found not approvable
- Sponsor advised to:
 - Compare intrapartum obstetrical practice between Sweden and US
 - Revise labeling and education program to bridge any gaps between practice styles
 - Conduct clinical validation study in US



27

Plan for US Study of STAN

- June 2003 Panel Meeting (closed session)
 - STAN Clinical Use Guidelines
 - Outline for 2-part study design
 - ❖ Part I: Non-interventional "Education Study"
 - ❖ Part II: Clinical Use Study



28

STAN - US Education Study

- 13 US readers
- Three different readings of 51 intrapartum tracings from Sweden
 - 1st before STAN training (FHR-only)
 - 2nd after STAN training (FHR-only)
 - 3rd after STAN training (FHR+ST)
- Comparison with 7 STAN experts



29

US Education Study Primary Hypothesis

- Mean correct decision rate among US raters is significantly greater when using FHR+ST data compared to using FHR-only data (3rd reading vs 2nd reading)
- Intervention "correct" for pH < 7.15
- Non-intervention "correct" for pH ≥ 7.15



30

US Education Study Secondary Hypotheses

- Mean correct decision rate among US raters *incorporating timing of intervention* is significantly greater when using FHR+ST data compared to using FHR-only data (3rd reading vs. 2nd reading)
 - gold standard for timing = median time of intervention of STAN experts
- Among intervention cases, difference in timing of intervention by experts and US raters < 20 minutes



31

Education Study Results

	FHR-only (2 nd exam)	FHR + ST (3 rd exam)
Correct Decision re Intervention	53%	69%
Correct Decision to Intervene + Timing	43%	59%



32

Education Study Intervention Rates by pH

	Cord Arterial pH		
	<7.05	7.06-7.14	>7.14
US raters (n=13)	90%	75%	39%
Experts (n=7)	90%	64%	9%



33

Rates of US and Expert Agreement to Intervene within 20 min by pH

	pH < 7.05 (N=9 cases)	pH < 7.15 (N=19 cases)
US and STAN experts	8/9 (89%)	14/19 (74%)



34

US Clinical Use Study

- Non-randomized, multi-center, 530 subjects
- 39 US investigators; 3 STAN experts
- Clinical management per STAN guidelines
- US investigators vs cord blood gases
- US investigators vs STAN expert consensus



35

Primary Endpoints

- Negative Predictive Value (NPV)
 - Probability that non-intervention results in a normal outcome in cohort of infants with NRFHR when STAN allows continued labor

NRFHR, STAN reassuring, no intervention, pH > 7.12
NRFHR, STAN reassuring, no intervention



36

Primary Endpoints

- Positive Percent Agreement (PPA)
 - Agreement of US clinicians with the majority decision of STAN experts on cause and time of decision to intervene when STAN Guidelines indicated intervention



37

Primary Endpoints

- Negative Percent Agreement (NPA)
 - Agreement between US investigator and STAN experts not to intervene when intervention is not warranted per STAN guidelines



38

Selected Secondary Endpoints

- Case-based analysis of all cases with pH <7.13
- Case-based analysis of cases in which intervention was indicated but US investigators did not intervene
- Operative intervention rates



39

US Clinical Use Study

- Education/Certification Phase
 - Self study
 - Onsite training (lecture, case discussions, equipment review and multiple choice test)
 - 42 investigators certified (3 excluded)



40

US Clinical Use Study

- Pilot Phase/Credentialing
 - Obtain STAN recording for minimum 5 cases
 - STAN data not used for pt management
 - Case discussions with site PI
 - Credentialing if STAN concepts adequately understood



41

US Clinical Use Study

- Pivotal Phase
 - 530 subjects monitored
 - 373 SVDs
 - 157 operative deliveries
 - ✦ 111 cesarean section
 - ✦ 46 operative vaginal deliveries
 - Cord artery acid base data on 88% (466/530)



42

Clinical Use Study Results

	Target	Outcome 95% CI	Met Hypothesis?
Negative predictive value (NPV)	75%	95.2% 91.2, 97.8	Yes
Positive % agreement (PPA)	75%	83.8% 68.0, 95.7	No
Negative % agreement (NPA)	75%	90.4% 87.8, 93.0	Yes



43

Positive Percent Agreement

- Result based on 31/37 agreement rate
- Different basis for intervention (1 case)
- US investigator did not intervene despite STAN indication to intervene in 5 cases
 - C-section unrelated to STAN guidelines (2)
 - Vacuum-assist vaginal delivery 50 minutes after expert recommendation (1)
 - SVD at 22 min and 28 min after expert recommendation (2)



44

Results - Selected Secondary Endpoints

- Case-based analysis of all cases with pH <7.15
- Case-based analysis of cases in which intervention was indicated but US investigators did not intervene
- Operative intervention rates



45

Safety Results

- 1 perinatal death 2 weeks postpartum during repair of hypoplastic left heart/aortic atresia
- 1 CVA at 28 hours
- 1 case of metabolic acidosis (pH 6.88; base def 14.4) 1 min Apgar 7; 5 min Apgar 8; not admitted to NICU



46

Comparison of Safety Outcomes in Swedish RCT and CUS

	Swedish RCT N=2519	Clinical US Study N=530
1 min Apgar <4	1.43%	2.08%
5 min Apgar <7	1.03%	1.13%
5 min Apgar < 4	0.08%	(no cases)
NICU	6.71%	6.98%
Encephalopathy	0.04%	(no cases)
Infant Death	0.08%	0.19%
Acidosis	0.69%	0.21%



47

STAN User Acceptability Results

STAN User Acceptability Category	User Response N=39
Adequate information throughout labor	88%
Easy to use	87%
Improved ability to assess fetus	68%
Influenced patient management	32%
Need to adjust sensors to improve signal	34%



48

STAN Training Program

	Swedish RCT	US Ed Study	US Clin Study	Proposed for Market Setting
Text & CD rom	X	X	X	X
Cert Test	X	X	X	X
Credentialing			X	X
Continuing Education				X



49

Summary

- Safety and effectiveness demonstrated in SRCT (April 2002 Panel)
- US version STAN Guidelines
- 2 US bridging studies
- 4-Component STAN education and training program



50

Neovanta STAN S31 Monitor

Statistical Design and Analysis of STAN Monitor Studies

Gene Pennello, Ph. D.

Diagnostics Branch

Division of Biostatistics

Center for Devices and Radiological Health



51

Recap of STAN Studies

Early Clinical Studies

sensitivity, specificity, PPV, NPV

Swedish RCT Clinical Outcomes Study

metabolic acidosis rate, intervention rate

US Bridging Studies

Education Study.
Can US clinicians be trained?

Clinical Use Study.
Can US clinicians apply training in clinical setting?



52

Outline of Talk

- Diagnostic Endpoints (se, sp, PPV, NPV)
- Education Study Findings
- Clinical Use Study Findings
- Summary



53

DIAGNOSTIC ENDPOINTS



54

Evaluating Diagnostics

Diagnostic Test Has Two Outcomes:

Test is + for condition of interest
Test is – for condition of interest

Evaluate trade-off between
falsely

detecting condition test + when condition present
detecting condition test + when condition absent

Uninformative test: tests + as often when condition is absent as when it is present.



55

Traditional Diagnostic Endpoints

Sensitivity, How often does condition
Specificity: forecast test result?

PPV[†], How often does test result
NPV[†]: forecast (diagnose) condition?

[†] Positive and Negative Predictive Value

Informative Test:

sensitivity + specificity > 100%
or PPV + NPV > 100%.



56

STAN Monitor

Condition of Interest:

Study	Cord artery blood pH	BD _{ecf} (mmol/L)
Early clinical studies	variable	variable
Swedish RCT	< 7.05	> 12
US Education Study	< 7.15	—
US Clinical Use Study	< 7.13	—

STAN Test Result Definition:

+ if clinician intervenes before SVD
– if clinician doesn't intervene before SVD



57

Study Design Limitation

Interventional Studies: *Once you intervene, the outcome for the case had you not intervened is unknown.*

NPV is unbiased. Proportion of non-interventions without condition.

Sensitivity, specificity, PPV are biased. Last outcome, at time of intervention, is carried forward (LOCF) to time of SVD.

CUS Study. Considers only NPV.
Education, Early Studies: LOCF bias.



58

EDUCATION STUDY (Question 1)



59

Study Design

51 Tracings in which STAN was Used:

Cord Artery pH Level		
< 7.05	7.05-7.14	≥ 7.15
9	10	32

“True” Intervention Status:

pH level < 7.15: intervene
pH level ≥ 7.15: do not intervene

NOTE: For interventions, pH level is an LOCF outcome.



60

Study Design

13 US Clinicians:

1. Read FHR-only (Exam 1).
2. Training and Certification.
3. Sequentially, read FHR-only (Exam 2), then FHR+ST (Exam 3).

7 EU STAN Experts:

1. Read FHR-only (Exam 2)
2. Read FHR+ST (Exam 3).



61

Endpoints

- **% Agreement w. True Intervention Status** mean over US raters¹
consensus of US raters*
- **% Agreement incl. Intervention Time**²
< ± 20 min. of median of STAN experts
- **Intervention Rate by pH level**²
→ sensitivity*, specificity*
- **% Agreement with STAN Experts***



¹ primary, ² secondary, * not pre-specified.

62

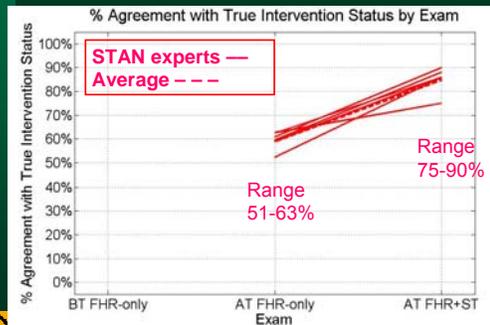
% Agreement with True Intervention Status

	Mean % Agreement Over Raters		Proportion of Raters Improving
	Exam 2: FHR-Only Read AT	Exam 3: FHR+ST Read AT	
All US Raters	53 %	69 %	12/13 (92%)
Certified Raters	55 %	69 %	9/10 (90%)
Rater 10 (Certified)	53 %	45 %	



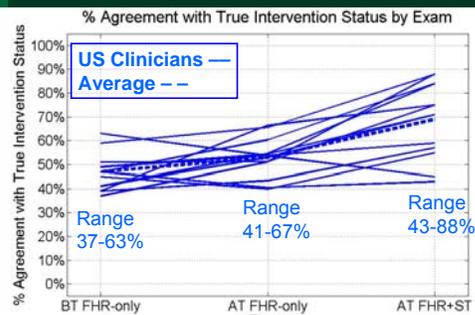
63

% Agreement with True Intervention Status



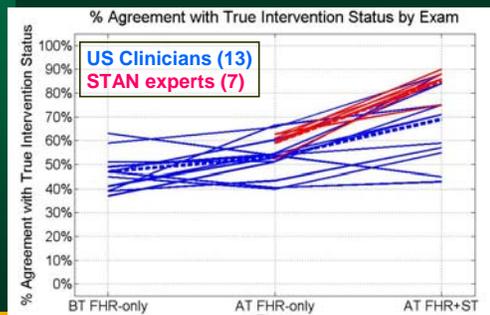
64

% Agreement with True Intervention Status



65

% Agreement with True Intervention Status



66

% Agreement with True Intervention Status

Primary Endpoint:

Mean % agreement over 13 US raters.

Primary Hypothesis:

Mean % agreement significantly > for Exam 3, AT FHR+ST read, than Exam 2, AT FHR-only read.



67

Test of Primary Hypothesis

Repeated Measures Analysis

	Exam 1: BT FHR-Only Read	Exam 2: AT FHR-Only Read	Exam 3: AT FHR+ST Read
Rater 1	51%	51%	88%
Rater 2	49%	55%	59%
...
Rater 13	39%	67%	75%

$p < 0.05$ for difference in mean % agreement between Exam 3 (69%) and Exam 2 (53%).

P value has not been validated yet.



68

Test of Primary Hypothesis

51*13*3 readings are not independent.

Three types of correlations are present:

Repeated readings over exam times.
Readings of 51 tracings by same reader.
Readings by 13 readers of same tracing.

Proper analysis accounts for all three correlations.

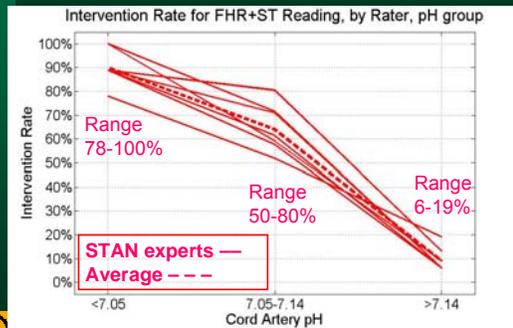
Repeated Measures Analysis

Accounts for 1st, 2nd correlations, but not 3rd.



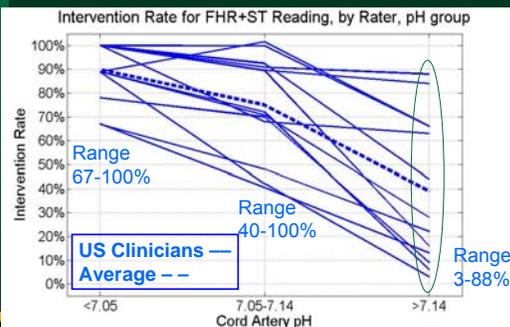
69

Intervention Rate for FHR+ST (Exam 3), by Rater, pH Group



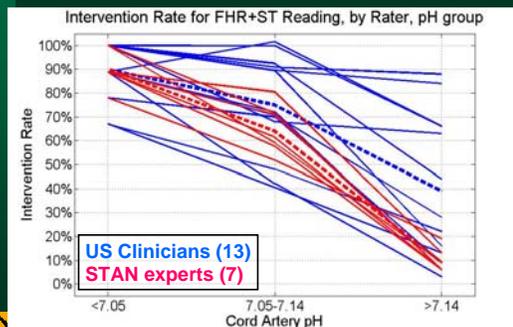
70

Intervention Rate for FHR+ST (Exam 3), by Rater, pH Group



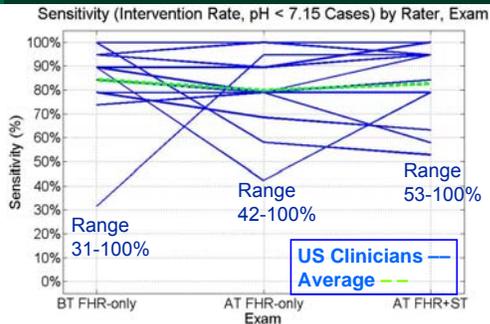
71

Intervention Rate for FHR+ST (Exam 3), by Rater, pH Group



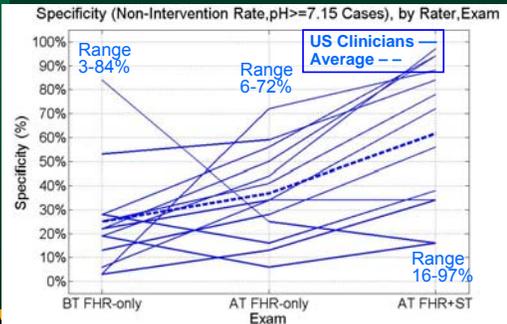
72

Sensitivity by Rater, Exam



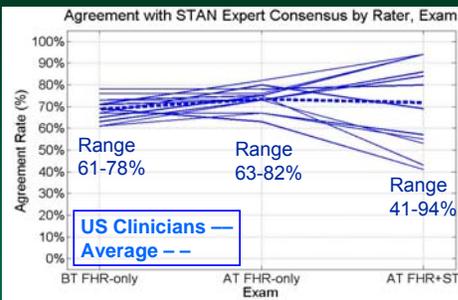
73

Specificity by Rater, Exam



74

Agreement w. STAN Expert Consensus by Rater, Exam



75

CLINICAL USE STUDY (Question 2)



76

Study Design

Non-Randomized, Single Arm Study
Cases Managed using FHR+ST
6 Centers
42 Investigators
530 Cases Enrolled
≥ 8 cases per investigator required



77

Primary Endpoints

Endpoint	Population	Hypothesis
Negative Predictive Value	Cases w. Non-Reassuring FHR tracings (N = 239)	NPV > 75%
Positive % Agreement [†]	All cases (N = 530)	PPA > 75%
Negative % Agreement [†]	All cases (N = 530)	NPA > 75%

[†] agreement with consensus of blinded, post-study readings by 3 STAN experts.



78

Primary Endpoint Target Values

NPV > 75%. Based on European data, NPV of 82% appeared achievable.

PPA > 75%. Based on Education Study, appeared to be achievable unless ST segment was ignored.

NPA > 75%. Based on Education Study, appeared to be achievable unless ST segment was ignored.



79

NPV Analysis (Sponsor)

Condition of Interest: pH < 7.13.

NPV Definition: Proportion of non-interventions for which pH \geq 7.13.

Sponsor Analysis

NPV 95.2% (180/189)
95% CI (91.2%, 97.8%)

Hypothesis NPV > 75% was met.



80

NPV Analysis (Sponsor)

Other cut-offs besides pH < 7.13 were used to define condition of interest.

For all cut-offs 7.10-7.15, hypothesis that NPV > 75% was still met.



81

NPV Analysis (Sponsor)

NPV 95.2% (180/189)

Denominator 189: Cases of NRFHR tracings for whom STAN allows continued labor. Includes non-interventions, *and interventions outside STAN guidelines.*

Numerator 180: & pH level \geq 7.13.



82

NPV Analysis (FDA Requested)

Exclude interventions outside STAN guidelines.

Rationale: They are interventions. NPV defined only on non-interventions. For interventions, outcome is missing.

NPV 95.5% (127/133),
95% CI (90.4%, 98.3%)

Hypothesis NPV > 75% still met.



83

PPA Analysis (Sponsor)

PPA: proportion of STAN expert interventions for which US clinicians (1) agreed to intervene, (2) agreed with timing of intervention.

PPA 83.8% (31/37)
95% CI (68.0%, 95.7%)

Hypothesis PPA > 75% was not met.



84

PPA Power Analysis

Sample size: 500 cases

Power: 80% to detect PPA > 75%

Assumptions for Power Calculation:

STAN expert intervention rate = 9%
(7% [37/528] observed)

True PPA = 90%, ignoring timing
(with timing 84% observed)



85

PPA Analysis (FDA Requested)

SVDs predefined as agreements with STAN expert decision to intervene if that decision was made within 20 minutes of delivery.

FDA request: Exclude such cases (17).

Rationale: Had SVD not been imminent, we don't know if US clinician would have intervened.

PPA 70.0% (14/20)
95% CI (45.7%, 88.1%)

Hypothesis PPA > 75% was NOT met.



86

NPA Analysis (Sponsor)

NPA: proportion of STAN expert non-interventions for which US clinicians agreed to not intervene

NPA 90.4% (444/491)
95% CI (87.8%, 93.0%)

Hypothesis NPA > 75% was met.



87

NPA Analysis (FDA Requested)

Interventions for failure to progress (FTP) considered as agreements with STAN expert decision to not intervene.

FDA request: Exclude such cases (82).

Rationale: Intervention is an intervention.

NPA 88.5% (362/409)
95% CI (85.0%, 91.4%)

Hypothesis NPA > 75% still met.



88

SUMMARY

Education Study (Question 1)

Certification test: 10/13 US Raters passed

Variation over raters in intervention rate.

Sensitivity maintained (on average)

Specificity improved (on average)

Primary hypothesis P value still under review. Mean % agreement with true intervention status improves from AT FHR-only read to AT FHR+ST read.

70% agreement w. STAN experts (ave.)



89

SUMMARY

Clinical Use Study (Question 2)

Primary Hypothesis NPV > 75% was met.

Primary Hypothesis NPA > 75% was met.

Primary Hypothesis PPA > 75% NOT met.

Conclusions robust to re-analysis.

PPA hypothesis not met probably because power was based on assumptions that were not precisely correct.



90