Author’s response to reviews

Title: Using fetal scalp stimulation with Doppler ultrasonography to enhance intermittent auscultation in low-resource settings: A diagnostic trial from Tanzania

Authors:

David Goodman (davidmgoodman@gmail.com)
Pendo Mlay (pendomlay1975@gmail.com)
Nathan Thielman (n.thielman@duke.edu)
Maria Small (maria.small@duke.edu)
John Schmitt (john.schmitt@duke.edu)

Version: 1 Date: 15 Dec 2018

Author’s response to reviews:

To the editor and reviewers:

Thank you for considering out manuscript for publication. We appreciate your thoughtful review. Your comments have helped to bring clarity to this study where it was lacking, and they have enhanced the paper. We continue to genuinely believe this study will help advance the field of low-cost fetal monitoring even though it is by no means a definitive trial that resolves all of the issues providers face around the world.

Please find our responses to your review below. We wish you all the best.

Sincerely,

David Goodman MD, MPH

Reviewer #1 (Jonathan Nelson MRCOG):

1. Whilst validation of the handheld Doppler devices with a CTG was nice to show a strong correlation, it should be acknowledged that there are no standards to say this degree of correlation means we are ok to procedure e.g. it was still your subjective choice as investigators to go ahead with this degree of correlation.
a. We changed the text to say “Based on these findings, we subjectively felt it was appropriate to proceed with the study.”

2. You mention using chi-square tests to compare groups, but should these be listed or mentioned in table 1? –

a. There are opinions based on various journals as to whether Table 1 should include p-values, particularly when the groups are not randomized. We chose not to include them, but have happily added them.

3. My main query or concern is surrounding how the study influenced intervention e.g. caesarean section. I understand caesarean rates were similar, but was there any change in management based in the IA+FSST group? I couldn't quite work out whether you were performing the FSST but then ignoring whether it was positive or negative and following previous protocols for intervention based on IA, or whether you were basing management on it e.g. timing of delivery.

a. We apologize for the confusion. We have changed the text in several places to clarify this and added Figure 1 to more thoroughly describe the protocol.

b. In the Setting and Practice section we added “The study protocol did not modify obstetric practice with respect to decisions for labor augmentation or cesarean delivery. Obstetric providers follow the WHO partogram and frequently use oxytocin to augment labor when a woman crosses the action line.”

c. The protocol was described more thoroughly in the methods section.

4. This links in with my other query as to the comment about this being unethical to perform currently as an RCT. If you were altering management in the IA+FSST group then I don't see how whether it was performed in this study design or RCT makes any difference because you are subjecting that IA+FSST group to a potentially unproven intervention. Assuming you weren't changing management based on the FSST then it might be worth acknowledging in the discussion that if this intervention were to be implemented it would need to involve quite widespread re-education. Without this it could result in increased intervention e.g. continued intervention in a group with standard 'fetal distress' but a normal FSST, and further intervention in a group without standard fetal distress with with an abnormal FSST.

a. We have clarified in the manuscript that we were not changing obstetric practice based on the FSST findings. Because we were collecting information we felt like we had to give some loose recommendations as described in the figure, but only one patient experienced 2 absent FSSTs spread out over one hour accompanied by non-reassuring fetal status and underwent cesarean delivery.

b. Such a clinical trial would necessitate strict adherence to a FSST-based protocol. It is possible that implementing FSST into labor protocols may lead to under-
diagnosis or over-diagnosis of non-reassuring fetal status, but our preliminary data indicate that an IA+FSST protocol would perform better. We believe that the addition of FSST to a protocol may help prevent cesarean deliveries because it gives providers a physiologically-based assessment of the long-term oxygenation status of the fetus. Significant re-education would be required should alterations to traditional IA be made. This would require in-service education distributed through the WHO and pre-service education in medical schools and midwifery schools in relevant areas around the world.

5. Also, whilst the increased examination in the IA+FSST group was mentioned, a concern in the resource-rich countries is the increased risk of infection/ascending chorioamnionitis with additional examinations, so at least this could be mentioned?

   a. We added this in the discussion: “It is possible that performing additional examinations may lead to an increased risk of intraamniotic infection, but this would need to be studied in future studies. The FSST was planned to be completed at times of routine vaginal exams when possible in order to minimize this risk. Investors may want to consider other methods such as vibroacoustic stimulation or Pawlik’s grip if they are concerned.”

6. I do feel the possible benefits of this sort of intervention could be explored further or referenced more as well. The potential neonatal morbidity/mortality through better identifying intra-partum hypoxia is mentioned, but the potential decrease in unnecessary caesarean sections is as big a positive in terms of maternal morbidity and future maternal/neonatal morbidity mortality. I wonder if this aspect should be mentioned more or at least be included as a measure in a future study?

   a. We stated “We believe that the addition of FSST to a protocol may help prevent cesarean deliveries because it gives providers a physiologically-based assessment of the long-term oxygenation status for the fetus.”

   b. It may be conjecture at this point to speculate further about the possibility for FSST to reduce cesarean deliveries.

Reviewer #2 (Andrew Weeks):

This is a relatively large study, with midwives using this technique in Africa for the first time. Low cost techniques for improving fetal monitoring during labour are understudied and potentially have huge impact on perinatal morbidity and mortality worldwide, which is currently unacceptably high. Furthermore, improving fetal monitoring during labour in LMIC has been highlighted as a research priority by experts. The authors' finding, that FSST enhanced intermittent auscultation improved the detection of significant acidaemia (pH< 7.0) from 27% to 70% is potentially of significance as a low cost, simple technique. I would therefore consider this topic area to be of great importance, and understudied.
MAJOR CONCERNS

1. It would be important to explain why cord gases were chosen for the diagnosis of "intrapartum fetal distress", whilst other important neonatal outcomes e.g. death, NICU admissions and APGARS were not recorded. Low pH does correlate well with poor outcomes long term, but the cut-off of 7.2 correlates very poorly with adverse neonatal outcome (mean arterial cord pH at birth is 7.24-7.26, so 7.2 is a very common outcome). A cut-off of 7.0 or 7.05 would be more appropriate and relate more to significant adverse outcomes.

   a. It is our view that umbilical artery blood gases collected immediately after birth is the standard for defining intrapartum fetal distress. This is consistent with the work that Steven Clark in the US has published. We recognize that this is an intermediate outcome, and therefore this current study should not serve as a definitive study that changes practice.

   b. We chose the pH<7.2 cut-off because this was the same cut-off used in the Clark and Rathore studies that examined the diagnostic utility of CTG and Pinard stethoscope to detect fetal scalp stimulation. We wanted this study to be consistent with those two studies so that the results would be easily comparable. We stratified the severity of acidosis and required the diagnosis of acidemia at birth to require an increased base deficit in order to be as stringent as possible.

2. Furthermore, cord pH readings are expensive (and so rarely available in low resource settings outside of research) and of limited clinical value in the management of newborns in this setting. Thus, a standard local method would have been more appropriate, especially given the ethical need to continue the supply of any research intervention after the end of the study (was this done?).

   a. We acknowledge that pH readings are expensive. The point of this study was to correlate scalp stimulation results with cord pH readings so that future studies and clinical practice would not necessitate cord pH readings. It was important at this stage in the development to correlate this method of fetal monitoring with physiological processes so that we could have confidence in planning a future study. If FSST had been found to not correlate well with cord pH then we would stop pursing it as a method for improving fetal monitoring in low-resource settings.

   b. Our manuscript stated in the conclusions that “Future studies should focus on neonatal outcomes rather than surrogate markers such as pH at birth.”

3. There are no details of consent methods for the woman - please could these be provided?

   a. Our “Ethical Review” section states: “All patients provided written consent prior to participation which is available upon request.”
b. We have changed it to read “All patients were approached by trained research nurses and counselled in Swahili. Written consent was provided prior to participation and is available upon request.”

4. The protocol for routine IA monitoring is not provided. In the IA and FSST group, how often were the women monitored if they had no risk factors? i.e. was the fetal heart only listened too once on admission and once when 9-10cm? The same detail is needed for the IA alone group.

a. We apologize that this was not made clearer in the initial manuscript. We added the following text to the methods section: “. Intermittent auscultation practice at KCMC follows WHO recommendations to auscultate the fetal heart rate every 30 minutes in the first stage of labor and every 5 minutes in the second stage of labor.”

b. Please also refer to the updated figure that shows the study protocol

MINOR CONCERNS

5. The abstract is not very clear about the methods and should be reworded to make clear the sequential process for recruitment of the patients.

a. We added a more thorough description to the abstract.

6. The abstract should mention that this was a pre- and post-intervention observational study.

a. The term “pre- and post-“ has been added in the methods section of the abstract

7. The setting of a referral centre in Tanzania is clear, but it would be useful to know the usual practice for fetal monitoring and if emergency CS is performed purely for fetal indications. The CS rate is 41% - how many of these are emergency CS?

a. Prior to the time of this study, KCMC did not delineate that in their data collection. Based on my broad inclusion criteria, it is reasonable to say that their emergency, primary cesarean rate is around 20%. Obstructed labor also lead to emergency (unscheduled) CSs.

b. We changed the text to read “The total cesarean section rate is 41%, which is divided almost evenly between emergency and repeat procedures”

8. The authors should provide information on the selection process for participants. What were the inclusion/exclusion criteria for the initial 50 patients? Were these the same as for the second study?
a. Thank you for pointing out that this was not clear

b. We changed the text under the Validation section to say “Prior to implementing FSST in the labour ward, it was imperative that we compare the ability of a handheld Doppler device to detect FSST with the gold standard CTG. Women were recruited if they had singleton, cephalic, term gestations and were undergoing fetal monitoring during active labor. Women were excluded if they had any risk factor requiring them to be a planned cesarean delivery such as placenta previa, previous myomectomy, active abruption upon arrival, eclampsia remote from delivery. The validation portion of this study took place in the first stage room with women that were in early labor and not on oxytocin so that the likelihood of fetal acidemia and therefore absent fetal scalp stimulations would be low.

9. Line 26 - typo - must "of" a means.

a. Thank you; this was changed to “labor providers must have a means”

10. The methods should be clear that the FSST is done vaginally. Some would do this using the 'Paulik's grip', so best to be clear in the text. It may also be worth mentioning this as an alternative in the discussion, given the invasive nature and infective risks of repeated vaginal examinations in labour.

a. In the Validation portion of the methods we state: Then midwives performed the FSST by stroking the fetal scalp with the tip of the examiner’s finger 5 times.

b. Then midwives performed the FSST by stroking the fetal scalp with the tip of the examiner’s finger 5 times. Transvaginal fetal scalp stimulation was deemed preferable to other methods of fetal stimulation such as clamping the scalp, vibroacoustic stimulation, or grasping the head with Pawlik's grip in a previous meta-analysis. [Skupski paper]

c. We added this in the discussion: “The marginal costs of adding 3-4 handheld Doppler units to a labour ward and performing an extended vaginal exam 1-3 times on each woman is negligible compared to the costs that would be incurred introducing continuous CTG. It is possible that performing additional examinations may lead to an increased risk of intraamniotic infection, but this would need to be studied in future studies. The FSST was planned to be completed at times of routine vaginal exams when possible in order to minimize this risk. Investors may want to consider other methods such as vibroacoustic stimulation or Pawlik’s grip if they are concerned.”

11. It is important to know how many high-risk patients there are in each group (e.g. PPROM, APH, postdates, pre-eclampsia, etc) in order to compare them, not just whether a CS was indicated. However, no details are given. Furthermore, there is no mention of
differing fetal monitoring for high and low risk patients, so we would have to assume that all are included and treated the same. Is this correct?

a. Yes, there was no difference in fetal monitoring for high-risk versus low-risk patients. That is why a distinction was not made in the manuscript.

b. We added a row for high-risk patients in Table 2 comparing the two groups. The IA+FSST group was higher risk on average, particularly including more cases of referrals for obstructed labor and fetal distress.

12. Only one obstetrician interpreted the CTG - it is well documented that there is often significant inter/intra observer error in CTG interpretation. Given that this is the Gold Standard in this study, can the authors provide detail on his / her experience or training, and what guidelines were followed for interpretation?

a. You are correct that CTG interpretation is fraught with difficulties and inconsistencies. The CTG strips were interpreted by David Goodman, a fellow of the American Congress of Obstetrics and Gynecology who as living in Tanzania during the project.

b. The CTGs were short ~10-minute tracings that were used to establish a baseline fetal heart rate and identify at 15 beat per minute acceleration lasting at least 15 seconds within 1 minute of the fetal scalp stimulation attempt.

13. It is not clear how soon after birth the gases were taken.

a. In the abstract we state that umbilical cord gases were collected immediately after birth. We changed this in the main text as well to state: Umbilical artery blood gases were collected immediately after cutting the cord following birth.

14. The pH outcome measure should be made clearer. It looks as if the main outcome was pH 7.2 with a BE of over -12. Is this correct? So if the pH was 6.9 with a BE of -11 this would not be an adverse outcome. Correct?

a. Yes, this is correct. A blood gas of that nature would indicate that a rapid-onset respiratory acidosis is present, but that it is not a significant metabolic acidosis that is capable of providing an hypoxic injury. This is based on the article: Ross MG, Gala R. Use of umbilical artery base excess: algorithm for the timing of hypoxic injury. Am J Obstet Gynecol. 2002 Jul;187(1):1–9.

i. First, we may assume that asphyxial injury does not occur until fetal base excess is ≤−12 mmol/L. Notably, most newborns with a base excess of ≤−12 mmol/L do not demonstrate neurologic injury. Even at levels of severe acidosis (base excess, ≤−16 mmol/L), most newborns either die or survive normally, with only a small proportion exhibiting cerebral palsy.
15. Why was audible decelerations under 110 chosen? Would baseline rate under 110 AND any decelerations auscultated be more in line with international guidelines?
   a. You are right, this was a typo during the editing. This should have said tachycardia, bradycardia, or audible decelerations.
   b. The text was changed to: Women received FSST hourly if they had the following risk factors: absent test on the initial exam, fetal tachycardia (>160 bpm), fetal bradycardia (<110 bpm), or audible fetal decelerations, meconium, or every 2 hours if they were receiving oxytocin for labor augmentation

16. Fig 1 - the standard methods for presenting diagnostic comparisons is in a 2x2 table. This would provide complete transparency of their data. The term 'fetal distress' is also not a good academic (or lay) term in my opinion. Better to use '34 with abnormal IA', etc which is less emotionally laden.
   a. Thank you for your opinion regarding the use of 2x2 tables. We felt that the current figure may be more appealing to readers, but will change to a 2x2 table if you feel that it is preferable.
   b. You are correct that fetal distress is no longer an appropriate term. The Brighton Collaboration Non-reassuring fetal status Working Group paper that significantly discouraged using the term fetal distress was released during this project. We did not modify our manuscript appropriately. We have changed fetal distress to non-reassuring fetal status and explained it more thoroughly in the text.

17. The process for fetal monitoring and intervention is not clear. I had assumed that woman were admitted, underwent a triage with FSST to either regular FSST or routine IA (maybe every 30-60 minutes?). Then if further FSST tests were abnormal, then the woman has a CS / operative vaginal birth. Is this correct?

18. Figure 1 presents in its legend (for the first time) that 'fetal distress' includes the presence of meconium-stained liquor. The grade of meconium for this definition should be provided as surely low-grade meconium is normal for a post-dates fetus? I would not have this as part of the diagnosis of 'fetal distress' but maybe just as an indication for FSST. I would also include accelerations and decelerations in this definition. Surely if decelerations were heard then this would be considered abnormal…?
   a. You bring up appropriate concerns. The decision to include meconium-stained liquor in the definition of fetal distress was three-fold. We chose to include it in the criteria for further monitoring because this is the way the Rathore study using a Pinard stethoscope to detect FSST was done, so we wanted to be consistent. We also anticipated that sensitivity for identifying fetal hypoxia would be low, so we wanted to include as many newborns as possible in our analysis. Also, the distinction between thin and thick meconium is more subjective than it simply
being present, so we felt that we could have a more consistent definition using these categories.

19. Given the increased use of vaginal examination, fetal infective outcomes should be provided.
   a. This would certainly be an important outcome to monitor in a future trial. We did not collect this data during this study.

20. At the end of the first CTG vs Doppler study, the authors should mention the outcomes for all patients - there is one missing in whom there was no agreement.
   a. We added the following clarification: “For the other test, the fetus had two decelerations noted on the CTG, but did have a present scalp stimulation test. Thirty minutes later the Doppler test was absent with a baseline of 130 and maximum response following stimulation of 138. This may represent deteriorating fetal status, but nevertheless the tests did not correlate. This particular mother eventually had a cesarean section with Apgars of 9 and 10 and no signs of adverse neonatal outcome.”

21. In table 1, the term 'referral' is unclear (are they referred in or out?). The term 'spontaneous labour' should be 'spontaneous onset of labour without need for augmentation'.
   a. In the Settings and Practice section we state that KCMC is an academic referral center. That would imply that we receive referrals, but we have clarified this in the text. We added this in the text: “Table 1 highlights small differences between the groups, notably the varying referral rates and spontaneous labor rates. More women in the IA+FSST group were referred from outside facilities.”
   b. We changed this to read “Of the women undergoing fetal monitoring in the IA group, 35% were experiencing spontaneous labor without the need for augmentation; for the remainder, labor was either induced or augmented with oxytocin intravenously.”

22. The 60% rate of augmentation of spontaneous onset labour requires comment, and a section in the methods to explain the reasons for use of augmentation.
   a. We added the following to the methods section: The study protocol did not modify obstetric practice with respect to decisions for labor augmentation or cesarean delivery. Obstetric providers follow the WHO partogram and frequently use oxytocin to augment labor when a woman crosses the action line.
   b. In the discussion we added: “Both cohorts had a significant proportion of women that received oxytocin augmentation in labor, 61% and 55% respectively for the IA and IA+FSST groups. Labor management was left to the discretion of the on-
call obstetric providers and followed the WHO partogram. It was common that oxytocin would be provided intravenously if a woman crossed the action line on the partogram. Augmenting labor is a fairly advanced obstetric practice that necessitates increased fetal monitoring beyond intermittent auscultation. Access to oxytocin combined with limited access to cesarean delivery has created a situation around the world that practice has progressed beyond the ability to safely monitor fetuses. It may be that adding fetal scalp stimulation to IA protocols would help make augmented labor safer for mothers and babies in areas where CTG is not available.”

23. In table 2 'asphyxia' is not defined.
   a. Thank you for pointing this out. We pulled this language directly from this study without defining it. As you mentioned, there is significant variation in how this term is used. We changed the text to say “intrapartum fetal hypoxia” in both places where birth asphyxia was used.

24. References 2 and 5 are the same.
   a. We apologize for this. We referenced two different Cochrane reviews and duplicated them in the bibliography. Your comment highlighted other inconsistencies in the bibliography that have been addressed.

25. There are differences in the number of recruits from the published protocol and the paper. The authors should explain the difference.
   a. We received free blood gas cartridges from Abbott Point of Care for this study. These were transferred from the US by one of the authors in two trips. Our supplies ran out faster than we predicted because the cartridges often malfunctioned. Our ability to enroll patients stopped when we ran out of cartridges.