PEER REVIEW HISTORY

BMJ Paediatrics Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Evaluation of a continuous Neonatal Temperature Monitor for low-resource settings: a device feasibility pilot study.
AUTHORS	Sosa Saenz, Sonia; Hardy, Mary; Heenan, Megan; Richards-Kortum, Rebecca; Dube, Queen; Kawaza, Kondwani; Oden, Maria

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Camilo Valderrama Institution and Country: Emory University Atlanta, GA, USA Competing interests: Biomedical Informatics Machine Learning Signal processing
REVIEW RETURNED	23-Feb-2020

GENERAL COMMENTS	Authors developed an affordable thermometer able to achieve comparable measurements to the gold standard used in high-income countries. The article was well written. However, I found the following minor issues: On page 9, letters used for Figure 4 do not correspond to those of the picture shown on page 21. In method sections, the authors did not explain completely how the measurement of NTM and the PM were compared. There are some details they gave in the results section that may be better on the method section since those details explain how the comparison was performed. The authors discussed some previous studies in the discussion section. However, in the introduction, the authors did not include any of those works in the background.

REVIEWER	Reviewer name: Katie Harron Institution and Country: UCL, UK Competing interests: no competing interested
REVIEW RETURNED	25-Feb-2020

GENERAL COMMENTS	How exactly did you determine that 35 patients were needed for this pilot study, and how were these sampled? Were there any differences in the laboratory test accuracies according to test (i.e. probe, temperature or monitor)? Please refer to the relevant appendix table within the text and provide explanations and titles for each of the appendix tables. How were the temperature measurements at admission taken? Table 1 – “average” – please be specific, is this mean?

	Figures 4A and B – what does “time” on the x axis refer to exactly? Why not display the whole time period? Figures4C-E, should the x axis be “temperature” rather than “temperature difference”? Please define what exactly is meant by bias (in the main text as well as a figure caption). Please check all labelling on these figures as there seems to be some confusion about the LOA and bias lines. How many different nurses took the temperature readings, and was there any variation in measurements by nurse? What were the different levels of experience of nurses? Were there any differences in categorisation of hypothermia according to type of monitor? Did you consider testing sensitivity and specificity for hypothermia to enable comparisons with other studies?

REVIEWER	Reviewer name: Daniele Trevvisanuto Institution and Country: University of Padova Competing interests: None
REVIEW RETURNED	27-Feb-2020

GENERAL COMMENTS	In this study, Sosa Saenz et coll. reported the results of a novel continuous temperature system assessed in a lab and in 39 neonates admitted to a neonatal ward in Malawi. Laboratory and clinical results show a good accuracy of this new system when compared to thermal standard methods and suggest that this new temperature monitor can be used in a low-resource setting. In addition, the authors report a good acceptance by parents and hospital staff. This is a well-written study that assess a worldwide problem in neonatology (hypothermia) and is of interest for pediatricians and healthcare staff taking care for baby in low-resource setting. Conclusions are consistent with the objective and the study results. Figures are clear and references updated. Limitations of the study include the limited number of enrolled patients in single center. I have the following suggestions: The system failed in 4 out of 39 (about 10%); may the authors add a comment on this result in the discussion? The authors reported that parents and staff were satisfied with the system, but did not report the method of evaluation (i.e a specific list of questions based on a Likert scale) neither the number of participants in this survey. Adding new technology in a new setting is sometime matter of concern. For example, the positioning of the belt was done by the research team or by the local staff? How can be possible to do the skin-to-skin (during KMC) if a device is placed on the abdomen of the baby? How does it remain fix?

VERSION 1 – AUTHOR RESPONSE

Comments from Reviewer 1

Comment 1: The article was well written. However, I found the following minor issues:

On page 9, letters used for Figure 4 do not correspond to those of the picture shown on page 21.

Response: We apologize for the mistake. This error has been corrected. The caption for Figure 4 and the description of Figure 4 in 'Results' have been updated to match each other.

In-text description: p.9 line 7 and p. 10 line 8

Figure caption: p. 16 line 31

Comment 2: In method sections, the authors did not explain completely how the measurement of NTM and the PM were compared. There are some details they gave in the results section that may be better on the method section since those details explain how the comparison was performed.

Response: As requested, we added some details from the results section to the methods section to explain the method of comparison on p. 8 line 8-11.

"Bland Altman plots, bias (mean difference), and 95% limits of agreement (LOA) were used to compare the agreement between the NTM device readings with the PM readings, as well as compare the NTM and PM readings with standard of care axillary temperature readings."

Comment 3: The authors discussed some previous studies in the discussion section. However, in the introduction, the authors did not include any of those works in the background.

Response: Thank you for this suggestion. We added additional information from these references into the introduction on page 4 line 19-21:

"Existing low-cost temperature monitoring devices that could be used in low-income countries, while sensitive enough to detect hypothermia, do not have the features necessary to guide clinicians in a hospital setting [11-13]"

Comments from Reviewer 2

Comment 1: How exactly did you determine that 35 patients were needed for this pilot study, and how were these sampled?

Response: We clarified how the sample size for the study was determined in the revised methods on page 7 line 10-13.

"A minimum of 35 subjects, with no less than 30% and no greater than 50% febrile subjects, were needed in accordance to ISO 80601-2-56 human subject population requirements. The standard requires at least 35 subjects in the group A1, 0 up to 3 months, for clinical accuracy validation[20]."

Comment 2: Were there any differences in the laboratory test accuracies according to test (i.e. probe, temperature or monitor)? Please refer to the relevant appendix table within the text and provide explanations and titles for each of the appendix tables.

Response: We added a sentence with regards to the laboratory testing vs clinical testing on page 10 line 11-13.

"The bias of the NTM vs PM in clinical testing is comparable to the average error between NTM and a reference thermometer observed during laboratory accuracy testing."

We also updated the supplemental files and included Supplemental Information file captions at the end of the manuscript on page 17 line 1-5. The Supplemental files are now cited in the text.

Comment 3: How were the temperature measurements at admission taken?

Response: We clarified in the revised methods that temperature upon admission was taken using axillary measurements (the current standard of care in the ward) on page 9 line 4 Table 1.

“Mean admission temperature (°C, Axillary with digital thermometer)”

Comment 4: Table 1 – “average” – please be specific, is this mean?

Response: Thank you for this suggestion. We updated Table 1, replacing “Average” with “Mean”. In addition, we changed “average error” to “mean error” on page 8 line 15.

Comment 5: Figures 4A and B – what does “time” on the x axis refer to exactly? Why not display the whole time period?

Response: Thank you for this suggestion. We clarified the meaning of the x-axis of said Figures on page 10 line 3-5.

“The x-axis Time (min) refers to the elapsed time of the temperature monitoring of a participant. A subset of the data is shown to highlight temperature tracking between both devices with significant resolution.”

Comment 6: Figures4C-E, should the x axis be “temperature” rather than “temperature difference”? Please define what exactly is meant by bias (in the main text as well as a figure caption). Please check all labelling on these figures as there seems to be some confusion about the LOA and bias lines.

Response: As suggested, we updated Figure 4C-E title, legend, and x-axis. The definition of bias was included in the revised methods section on p. 8 line 9.

“Bland Altman plots, bias (mean difference), and 95% confidence interval (CI) limits of agreement (LOA) were used to compare the agreement between the NTM device readings with the PM readings, as well as compare the NTM and PM readings with standard of care axillary temperature readings.”

Comment 7: How many different nurses took the temperature readings, and was there any variation in measurements by nurse? What were the different levels of experience of nurses?

Response: We clarified the method of temperature reading and setup in the text on page 12 line 21-23.

“During this pilot study, all SOC measurements and clinical set up were performed by one trained study nurse. Further research is needed to determine whether the device can be used consistently across users.”

Comment 8: Were there any differences in categorisation of hypothermia according to type of monitor? Did you consider testing sensitivity and specificity for hypothermia to enable comparisons with other studies?

Response: To assess the potential for differences in temperature classification, as an alternative to calculating sensitivity and specificity, we used Bland-Altman analysis to characterize the limits of agreement between the two monitors. This is an accepted standard to validate accuracy of temperature monitors. Future work will focus on the clinical performance (sensitivity and specificity) for detection of hypothermia under routine clinical use conditions.

Comments from Reviewer 3

Comment 1: The system failed in 4 out of 39 (about 10%); may the authors add a comment on this result in the discussion?

Response: We commented on the system issues in the discussion section on page 10-11 line 24-26 and 1-2.

“Out of the four system issues reported, two were problems with the interface between the data collection software and the NTM device, one was a connection issue from temperature probe to monitor. On one instance, the belt slipped down the abdomen, resulting in a gap between the temperature probe and the infant. These issues will be addressed in future iterations of the NTM device.”

Comment 2: The authors reported that parents and staff were satisfied with the system, but did not report the method of evaluation (i.e a specific list of questions based on a Likert scale) neither the number of participants in this survey. Adding new technology in a new setting is sometime matter of concern. For example, the positioning of the belt was done by the research team or by the local staff?

Response: Thank you for this comment. We clarified in text that the feedback was based on informal conversations with clinical staff on page 11 line 3-4.

“In informal conversations with clinical staff, they reported the NTM device easy-to-use and interpret…”

Comment 3: How can be possible to do the skin-to-skin (during KMC) if a device is placed on the abdomen of the baby? How does it remain fix?

Response: We added a sentence in the discussion about the use of the NTM device during KMC in a hospital setting on page 12 line 7-9.

“Furthermore, NTM’s portability could also allow for continuous temperature monitoring during KMC in a hospital setting. The belt would secure the temperature probe on the baby’s abdomen in the same way as in the NICU, and the monitor could be placed by the mother’s bedside.”

We added a sentence in the conclusion to encompass the feedback from the reviewers to take into consideration for further studies on page 13 line 1-3.

“The use of NTM in the KMC ward, NTM impact on outcomes and nursing care within the NICU, and the accuracy of the device when placed by different users should be explored in further studies.”