Author’s response to reviews

Title: Community Paramedic Point of Care Testing: Validity and Usability of Two Commercially Available Devices

Authors:

Ian Blanchard (ian.blanchard@ahs.ca)
Ryan Kozicky (ryan.kozicky@ahs.ca)
Dana Dalgarno (dana.dalgarno@ahs.ca)
Justin Simms (justin.simms4@gmail.com)
Stacy Goulder (stacy.goulder@ahs.ca)
Tyler Williamson (tyler.williamson@ucalgary.ca)
Susan Biesbroek (Susan.Biesbroek@ahs.ca)
Lenore Page (lenore.page@ahs.ca)
Karen Leaman (Karen.Leaman@cls.ab.ca)
Suzanne Snozyk (Suzanne.Snozyk@cls.ab.ca)
Lyle Redman (Lyle.Redman@cls.ab.ca)
Keith Spackman (Keith.Spackman@ahs.ca)
Christopher Doig (chip.doig@gmail.com)
Eddy Lang (eddy.lang@ahs.ca)
Gerald Lazarenko (gerald.lazarenko@ahs.ca)

Version: 1 Date: 23 Jan 2019

Author’s response to reviews:

2019 01 09
Dear Dr. Tu,

Thank you for your suggestions regarding our manuscript entitled “Community Paramedic Point of Care Testing: Validity and Usability of Two Commercially Available Devices” (EMMD-D-18-00094). The manuscript has now been revised with changes highlighted in yellow below and tracked in the revised manuscript.

We hope that you will find these revisions acceptable and we look forward to your response.

Sincerely,

Ian Blanchard

Reviewer 1 comments:

Thank you Dr. Mikkelsen for your comments and suggestions.

1a. The concept of community paramedics is not uniformly applied in Europe. The authors might elucidate the function of said group of professionals. Thank you for this suggestion, we have provided additional information for the reader in the Introduction and methods sections in the tracked changes version of the manuscript. Specifically we have added in the intro:

CP’s receive training in addition to their formative training, and generally have a broader scope of practice compared with regular duty paramedics.

And in the methods:

In addition to their formative paramedic training, CPs receive 21 days of additional training on assessment and treatment skills specific to this patient population. These skills consist of geriatric medicine, social determinants of health, advanced cardiopulmonary assessments, additional pharmacology (in particular antibiotics), urinary catheterization, wound closure (sutures, staples, adhesives), accessing central venous access devices, specimen collection, which includes wound, throat, and nasopharyngeal swabs, and blood and urine specimens.

1b. The study is lacking slightly in description of chain of events. It is thus unclear which criteria were applied for enrolling patients? The authors state that 1,649 events were screened for
enrollment, 174 had a blood draw, with 108 events (62.1%) enrolled from 73 participants. The enrolling of patients is incomprehensible: 1649 patients were screened. This part is understandable. However, the next part is incomprehensible: In 174 patient care events, a blood sample was drawn. How do the authors end up with 108 events (what is an event) in 73 patients (what about the remaining 101 "events"? What did they represent?) I realize that Figure 1 obviously helps. However, the number of patients - and the number of blood samples and the number of failed blood samples is not crystal clear.

Thank you for identifying this error. The issue was the inattention to the description of the unit of measure (event and patient) in Figure 1. Specifically we have changed “Patient Care Event” for “Episode of Care” as we propose this may better describe for the reader what is meant. We have also made changes and additions to the methods and results sections, and aligned figure 1 (see 2019 01 09 Figure 1 Revision 1), which we hope will clarify the enrollment for the reader. The text in the methods and results now reads respectively:

Patients were not excluded if they already had been consented into the study. In other words, one patient may have been enrolled multiple times in the study if they had multiple episodes of care that required a blood draw.

Of 1,649 episodes of care screened for enrollment, 174 episodes of care had a blood draw, with 108 episodes of care enrolled in the study from 73 participants (Figure 1).

1c. Why such a discrepancy between enrolled patients and number of events? Did the blood sampling fail in 37.9% of the cases? Thank you for this question. Like many CP programs, this program often is called to a single patient multiple times. In other words we may have patients that have multiple episodes of care. Patients who had already been enrolled in the study, were not excluded if they presented with a second episode of care. Out of 174 episodes of care that involved a blood draw, 108 (62%) were enrolled. Of those episodes that were not enrolled, the primary reasons are listed in Figure 1 under the major headings of “Episodes of Care with Blood Draw, but did not meet inclusion criteria” and “Episodes of Care with Blood Draw that met inclusion, but were not enrolled”. We have added percentages into the Figure to highlight for the reader the proportion meeting a characteristic. Of the 66 episodes of care that were not included, 50/66 (76%) were because the patient did not meet inclusion, an analyte was ordered that was not included in the study, or the patient declined consent (please see 2019 01 09 Figure 1 Revision 1). We hope that the addition of the percentages to describe proportion in Figure 1 should clarify this for the reader.
1d. Should one apply the intention-to-treat principles, based on my understanding of the reported results, the results are rather less favorable and claimed in the abstract. I am not sure that the claim that CP programs can expect valid results is supported by the data presented. Thank you for this observation. In terms of the intention-to-treat principle, we would propose it is not appropriate here as no patients or episodes of care were randomized to receive an intervention. All patients, and all of the patient’s episodes of care were eligible for inclusion. While 48% of episodes were not included in the study, 76% of these episodes were because the patient did not meet inclusion, an analyte was ordered that was not included in the study, or the patient declined consent. We would suggest therefore that the exclusion of eligible episodes of care were less than 10% (16 of 174).

In terms of the results not supporting our conclusions, we would propose that the results support that CP programs can expect clinically valid readings. This is based on:

1. A low number of critical discrepancies (less than 1%)
2. That POCT reported critical values when the laboratory did not, which would have resulted in over-scrutinizing patients

To emphasize that we propose these results are clinically valid, we have add “clinically” to both the abstract and main body conclusions.

1e. It is interesting that the i-Stat quality control seems to fail: The proportion of out-of-range results by device number 6 was (significantly) larger than in the other i-Stats. Do the authors have any supplemental comments? Did the company offer any explanation?

Thanks for this question. As outlined in the manuscript, we thoroughly investigated this issue and do not have a ready explanation for these results. In discussion with the manufacturer they felt it was most likely associated with the preparation of the sample by the paramedics involved. While other paramedics that used this device used other devices, it may have been the two primary paramedics that used this device and their technique in sample preparation. We simply don’t know for certain the cause of this variation.

We hope that our current wording in the discussion and limitations section is sufficient for readers of the article:
Discussion

Lines 536-44

When individual devices by manufacturer were compared, one i-STAT® device accounted for over half of all out-of-range i-STAT® results. This device did not have unusual incidents logged, nor was it exposed to extreme temperatures. All weekly QC testing was completed on the device. The cartridge lots were used by other i-STAT® devices and hence not unique to this device. Four CPs used this device, with two of the CPs using other i-STAT® devices in the study. It is therefore unknown why this particular device would return more out-of-range results compared with the other i-STAT® devices. For epoc®, none of the devices had unusually large numbers of out of–range results compared with each other, although two CPs using one device collected 12 of the 17 out-of-range Hct results on epoc®.

Lines 546-7

It is difficult to determine retrospectively what may have been the cause.

Lines 561-2

The results underscore the importance of proper training, initial device validation, daily and weekly QC checks, split sample testing, and handling and care of POCT devices.

Lines 592-4

Field observations of POCT use were unable to be conducted during this study, but would provide important information on why the device error rates changed between QC and patient testing and some of the outliers found in the comparison of the POCT with laboratory tests.

Limitations

Lines 625-9

While all attempts were made to train and assess competence in CPs involved in this study, no observational quality assurance was performed to ensure good technique by CPs while out in the field. Differences in technique associated with mixing and storage may have affected individual samples, and individual differences in how a CP tested a sample may explain some of the observed device discrepancy.

If. A problem on line 435 that should be solved.

Thank you for this. Issue resolved.
Reviewer 2 comments:

Thank you Dr. Russell for your comments and suggestions.

2a. Line 164: in the UK we say "took blood" rather than "had a blood draw".

Thanks for this suggestion. Based on previous suggestions we made a number of changes that removed some of this wording. It now reads:

Of 1,649 episodes of care screened for enrollment, 174 required a blood draw, with 108 episodes (62.1%) enrolled from 73 participants.

2b. Line 167 & several other places: Shakespeare used "compare A to B" to mean "liken A to B" & "compare A with B" to mean "contrast A with B". British books advising on English usage support this useful distinction. Please consider adopting it.

Thanks for the suggestion. As in Medicine, scholars in English grammar often disagree, and sometimes their arguments don’t reflect the practical use of language. One of the authors is more familiar with Strunk and White’s The Elements of Style. His English Professor mother was dismayed that when he finished University, she thought he had failed to learn one language. That author has been sacked and will not provide further editing. We have made multiple changes throughout the manuscript to reflect your advice.

2c. Lines 173 & 174: rather than "fewer field blood analysis device-logged errors occurred in i-STAT compared to epoc", I offer "fewer errors in analysing blood in the field occurred with i-STAT than with epoc".

Thanks for this suggestion. The text has been changed. It now reads:

There were no statistically significant differences in device logged errors between i-STAT® and epoc® (p=0.063).

We would propose that your suggestion changes the meaning slightly from errors that were logged by the device to errors in analyzing the blood. Some logged errors were not from user error, but may have been from equipment issues.

2d. Lines 174 & 175 & several other places: in comparing i-STAT with epoc, it will help readers to follow the "Consolidated Standards Of Reporting Trials (ConSORT)" by reporting a confidence interval (CI) for the difference in error rates that includes zero rather than leave them to combine separate CIs for the 2 rates.
Thanks for this suggestion. We agree, providing a confidence interval for the difference between the two devices would allow the reader to directly compare the two devices easier. Our rationale for providing the two device specific estimates is to allow the reader to quantify the estimated proportion of critical or out of range results with a precision estimate for each device.

In both instances an estimate of the error with a 95% CI is likely more useful than the difference between the two devices (which does not allow the reader to get a sense of the device specific error rate, only the difference in device specific error rates). We considered also including the difference with 95% CI’s, but felt that this made the text difficult to read as we were presenting three 95% CI’s in rapid succession.

We are open to adding the difference between the devices, but would advocate for keeping the device specific error rates with 95% CI.

2e. Lines 176 & 177: as these Conclusions will attract most readers, they need to be more friendly to them.

Thank you for this comment. We have made some changes to the conclusions, it now reads:

CP programmes can expect clinically valid results from POCT. Device usability assessments should be considered with any local implementation as the two POCT systems have different strengths.

2f. Line 221 & other places: please avoid using the forward slash as an all-purpose conjunction & say "patients or results".

Thank you for this suggestion. All forward slashes have been removed and the appropriate conjunction added.

2g. Line 238: please omit ", in terms of validity," which confuses this otherwise clear sentence.
Thank you for this suggestion. The sentence now reads:

This suggests that these two devices are functionally similar in the hands of laboratory personnel in a controlled environment.

2h. Line 261 & 264 & other places: please use "program" for computer programs & "programme" for programmes of work.

Thank you for this suggestion. We have made numerous changes throughout the text to ensure consistency in the use of the word programme. As the Canadian Safety and Security Program is a proper name we have left it unchanged.

2i. Line 264: does "using six non-transport capable CP units" mean "using six vehicles that cannot convey patients"?

Thank you for this suggestion. We have clarified this statement and found that there was indeed one vehicle that could transport patients. The sentence now reads:

At the time of this study there were 19 active CPs in the programme using five vehicles that cannot convey patients and one that can.

2j. Line 277: please substitute "On scene" for "While on-scene", thus avoiding both attaching that phrase to "a portion of blood" & the superfluous hyphen.

Thank you for this suggestion, the sentence now reads:

On scene a portion of the drawn blood was also used for POCT testing (split-sample).

2k. Lines 315, 332 and other places: please remove more superfluous hyphens.

Thank you for this suggestion. We have removed multiple superfluous hyphens.

2l. Lines 316 to 320: please rewrite this indigestible sentence.

Thank you for this suggestion. The sentence now reads:
The curriculum included training on the operation of i-STAT® and epoc® devices and troubleshooting strategies. CPs also received an overview of the research study, ethics, consent procedures, additional equipment, documentation and data collection.

2m. Lines 340 to 342: please elaborate these condensed methods.

Thank you for this suggestion. We are not sure what further elaboration is required? We would propose that the sentence has included the targeted sample, the margin of error, and the assumptions in both prevalence of out of range results and device sensitivity. Are you looking for further justification of the assumptions in the sample size calculation?

2n. Lines 345 to 347: British version of Good Clinical Practice strongly recommends avoiding Excel in randomised trials (which includes this study) as it generates no audit trail. Nevertheless I should be willing to publish this useful paper provided the authors advise readers to avoid this methodological weakness in Limitations.

Thank you for this suggestion. We have added the following to the limitations section and hope it will suffice:

Finally, data were entered into a Microsoft Excel spreadsheet, which increased the risk of unplanned alterations and unwelcomed access to the data. To mitigate this limitation the spreadsheet was password protected, kept on a limited access shared drive, and only two research associates had access to the file.

2o. Lines 349 & 350: as the Normal distribution is robust to small departures, it would be more usual to say "medians & inter‐quartile ranges for data that clearly diverge from normality".

Thank you for this suggestion. The sentence now reads:

Descriptive data are reported as means and standard deviations for normally distributed data, or medians and inter‐quartile ranges for data that clearly diverge from normality.

2p. Line 430: in accordance with ConSORT (also mentioned in comment 4) please add confidence interval.

Thank you for this suggestion. 95% CI added. The line now reads:
Participants scored the i-STAT device a mean 24.4 points higher (95% CI 6.9, 42.0) than the epoc device.

2q. Lines 469 & 470: this sentence reports a Result rather than a point for Discussion.

Thank you for this comment. In the Results section we explicitly mention these results and further outline them in Table 3. Our rationale for briefly mentioning it again in the Discussion section is to provide context to the discussion, which focuses on possible explanations for the results found.

2r. Line 482: what is "cold-chain" please?

Thank you for this question. We have removed the term cold chain, the sentence now reads:

The cards or cartridges were likewise not exposed to any known extreme temperatures.

2s. Line 572 to 574: not many readers will understand this sentence.

Thanks for this feedback. We have changed the last sentence of the conclusions to now read:

Device usability assessments should be considered with any local implementation as the two POCT systems have different strengths

2t. Line 593: please write "calculation verification" in full throughout!

Thank you for this suggestion; changed throughout document.

2u. Line 630: please use BMC reference style, which does not include issue number, throughout.

Thank you; issue number removed.