Author's response to reviews

Title: A Prospective Study To Evaluate The Accuracy Of Pulse Power Analysis To Monitor Cardiac Output In Critically Ill Patients.

Authors:
Maurizio Cecconi (maurizioceconi@hotmail.com)
Jayne Fawcett (jayne.fawcett@stgeorges.nhs.uk)
Michael Grounds (Michael.Grounds@stgeorges.nhs.uk)
Andrew Rhodes (andyr@sgul.ac.uk)

Version: 4 Date: 22 August 2007

Author's response to reviews: see over
Author’s response to reviews

Title: A Prospective Study To Evaluate The Accuracy Of Pulse Power Analysis To Monitor Cardiac Output In Critically Ill Patients.

Authors:
Maurizio Cecconi (maurizioceccconi@hotmail.com)
Jayne Fawcett (jayne.fawcett@stgeorge.nhs.uk)
R Michael Grounds (Michael.Grounds@stgeorges.nhs.uk)
Andrew Rhodes (andyr@sgul.ac.uk)

Author’s response to reviews:
We would like to thank the reviewers for their time and commitment to improving this paper, which they undoubtedly have. We have provided a point by point response to the reviewers comments together with a clean copy of the revised manuscript.

Reviewer 1 (Mervyn Singer)

1. Abstract:
   How can the median value be 5 l/min yet the IQR be 5.1-9 L/Min?? These data also contradict the data shown in the Results section (page 7, 2nd paragraph).

   We thank you for pointing this out to us. The reviewer is quite correct. We have corrected the manuscript accordingly.

2. ‘The cardiac output remains acceptable for up to 4 hrs in patients with haemodynamic instability ‘ – what evidence do you have that the patients are ‘unstable’? I note that they were (page 5) ‘deemed so by their treating clinicians’ but this is both vague and highly subjective. On the other hand (page 6) measurements at baseline, 1, 2, 4 and 8 hrs were performed only during intervals relatively free of haemodynamic change yet these were achieved in virtually all cases. On page 7 (Results) it is stated that only 6 of the 14 patients had a cardiac output change >15% from baseline – 15% does not constitute a particularly large change and 8 of the 14 had an even smaller change. Either define instability more stringently or remove it altogether.

   We have changed the manuscript in all of these places to agree with what the reviewer has said. We now no longer describe this population as being haemodynamically unstable, which we were unable to accurately and robustly define. They are now described as a population of critically ill patients who required cardiac output monitoring.

   ‘Adult critically ill patients who for clinical reasons were being monitored with the LiDCOplus cardiac output monitor on the General Intensive Care Unit at St George’s Hospital were enrolled into the study.’
3. Results (page 7, 2nd para): It is rather misleading to talk of ‘excellent correlation’ yet in the following sentence to highlight the large error at 1, 2, 4 and 8 hours. The correlation is only excellent due to the single outlying point of 19 l/min. The text should be corrected.

We have changed the text in accordance with the reviewers wishes.

‘Data for PulseCO and LiDCO at 1, 2, 4 and 8 hours were significantly correlated with \( r^2 \) values being greater than 0.86 at all time points [Figure 1]. This correlation was associated with acceptable levels of bias and limits of agreement for the first four hours of the study, however at eight hours following calibration the PulseCO device had a percentage error that was outside of the acceptable range.’

Reviewer 2 (Nicholas Linton)

A casual reader might look at figure 1a and erroneously conclude that there is evidence that PulseCO is useful within one hour of calibration - the differences between the methods are relatively small (right plot) in comparison to the large range of cardiac outputs (left plot). However, the range of cardiac outputs in fig 1a might be due to the large range of cardiac outputs at calibration. The calibration of PulseCO with LiDCO introduces a confounding variable (the calibration value): the apparent agreement between PulseCO and LiDCO in figure 1 is probably spurious.

The authors state that the device “is sold as a monitor of cardiac output and this is what we tested.” My understanding from the title of the paper is that they set out to test the Pulse Power Analysis component of the device. This is designed to track changes in cardiac output following calibration, and trends of cardiac output are displayed on a screen. If the authors actually set out to test the combined accuracy of Pulse Power Analysis including the LiDCO calibration, then they should have used an independent reference method such as thermodilution.

We agree with all of these points that the reviewer has made. This is indeed a recognized weakness of doing a linear regression analysis for this type of study. We have added the following statement to the discussion to adhere with these views.

‘This agreement may in part relate to the wide range of cardiac outputs that our patients presented with in comparison to the relatively small changes in the variable seen over the first few hours. This raises the possibility that the relationship may in fact be a spurious artefact of the analysis rather than a real phenomenon.’
We also agree that what we have studied is the performance of the pulse power algorithm and not the absolute accuracy of the whole device (LiDCO and PulseCO). This is in accordance with what we initially set out to do. If we had wanted to study the whole technology we would have used an independent reference, however as we were mainly interested in the performance of the PulseCO we decided to use a validated independent technique (LiDCO). We have amended the manuscript in a number of places to clarify this situation.

‘This study has assessed the ability of the pulse power algorithm to maintain its accuracy when compared to the LiDCO against time. It has not compared the accuracy of the combination of LiDCO and PulseCO to give the absolute cardiac output. This would need a further independent measurement technique of cardiac output as a reference.’

The authors also state that the analysis that I have suggested is not possible because the exact precision of LiDCO at determining changes in CO is not known. This does not prevent the analysis. Previous publications have included analyses of changes in cardiac output. I maintain that if there is no agreement between “changes in PulseCO” and “changes in LiDCO” then the authors cannot conclude their study with a statement that implies that the pulse power algorithm is acceptable “for up to four hours in patients with haemodynamic instability”.

We have re-analyzed the data exactly as the reviewer has suggested. We have now included the percentage changes of cardiac output as measured by PulseCO when compared to similar changes in LiDCO. The results are included in the results section and also as Figure 2 (additional). We thank the reviewer for insisting we did this as it has strengthened the paper.

1. I agree that a recommendation based upon time is necessary but I am not convinced that it is sufficient. Is PulseCO accurate for four hours?

This has been answered in response to the points above. Changes in the PulseCO have a significant correlation with similar changes in LiDCO over the first four hours. There are also good agreements between the absolute numbers seen on Bland Altmann analysis and percentage errors as described by Critchley and Critchley. We agree that none of these analyses are foolproof, however they are all completely in accordance with the methodologies described by Balnd and Altmann and they are also the same approaches as are commonly used in many similar papers.

2. See above.

3. The authors say that they have identified a problem with the PulseCO system. I agree. My concern is that there is no foundation to their conclusion that “the pulse power algorithm … remains acceptable for up to four hours in patients with haemodynamic instability”. They have not shown that PulseCO can track
changes in cardiac output. Most of the patients in this study only had changes in cardiac output of less than 15%.

This analysis has now been included in the paper as asked for.

4. Reference 7 is now an article about horses and still contains no information about the Pulse Power Algorithm. Please quote a reference that describes the Pulse Power Algorithm.

This reference has been removed and the manuscript appropriately changed in the corresponding parts.

5. The authors say that their wording is correct. “Bias” is defined as the “mean of the differences” between the two methods (see Bland and Altman). How did the authors calculate the “standard deviation of the bias” (i.e. the standard deviation of the mean of the differences)?

We calculated the bias exactly as described by Bland and Altmann. The analysis was performed with Graphpad Prism software as now described in the text. On reflection we think that the reviewers wording is more likely to be correct and we have amended the manuscript accordingly.

‘This error rate was quantified by dividing twice the standard deviation of the differences by the mean cardiac output for both techniques.’

6. I think it would be useful to mention the response to phenylephrine because this is a situation where accuracy would have been clinically useful.

Unfortunately we do not have the data to describe the response to phenylephrine. We agree with the reviewer that it would have been of interest, however as we did not perform this study have not included this in the text.