Author's response to reviews

Title: Predictors of the accuracy of pulse-contour cardiac index and suggestion of a calibration-index: a prospective evaluation and validation study.

Authors:

Wolfgang Huber (Wolfgang.Huber@lrz.tum.de)
Jonas Koenig (koenig.jonas@hotmail.de)
Sebastian Mair (Sebastian.Mair@tum.de)
Tibor Schuster (tibor.schuster@mail.mcgill.ca)
Florian Eyer (Florian.Eyer@lrz.tum.de)
Veit Phillip (Veit.Phillip@lrz.tum.de)
Caroline Schultheiss (Caroline.Schultheiss@lrz.tum.de)
Philipp Thies (Philipp.Thies@lrz.tum.de)
Ulrich Mayr (Uli.Mayr@lrz.tum.de)
Henrik Einwächter (Henrik.Einwaechter@lrz.tum.de)
Josef Hoellthaler (josef.hoellthaler@siemens.com)
Roland M Schmid (Roland.Schmid@lrz.tum.de)
Bernd Saugel (Bernd.Saugel@lrz.tum.de)
Matthias Treiber (Matthias.Treiber@lrz.tum.de)

Version: 2
Date: 26 March 2014

Author's response to reviews: see over
Dear Editors,

Thank you very much for reviewing our manuscript entitled "Predictors of the accuracy of pulse-contour cardiac index and suggestion of a calibration-index: a prospective evaluation and validation study" (Critical Care manuscript: 1782168900103678).

We are very grateful for your interest and the opportunity to transfer a revised version of the manuscript to BMC Anesthesiology.

Therefore, we tried to consider the reviewers' concerns and suggestions and hope that we were able to respond sufficiently to their comments. We appreciate the reviewers' valuable comments and suggestions that really help to improve our manuscript.

You will find a point-to-point reply to the reviewers' comments and a detailed discussion with reference to the reviewers' concerns in this letter of response.

In addition to a “clean copy” of the revised manuscript you will find a revised "tracked changes" version (marked-up revision) of the manuscript with changes highlighted.

We hope that you find our revised manuscript worth to be published in BMC Anesthesiology.

Thank you very much in advance,

Yours sincerely

Wolfgang Huber, MD
Point-to-point-reply to reviewers’ comments and discussion regarding reviewers’ concerns:

Ad Reviewer #1:

5. The only thing I am missing is that emphasis is lacking on the actual “numerical message” of the results. Neither in the “Practical implications”, “Conclusions”, nor in the “Key messages” it is stated that what figures, ie.: what percentage changes of Clpc, should the clinicians look at and take seriously as indicators for recalibration. It is not easy to fish out these data from the paper, therefore, adding these to practical implications would definitely enhance the message to the reader.

Answer: This is a very good suggestion really improving the clarity of our paper. Indeed, the multiple analyses needed to be summarized by a pragmatic approach that can be readily applied by the physician irrespective of future improvements of algorithms etc.. Therefore, we included a statement in the practical implications paragraph, “summarizing different analyses of this study, re-calibration should be considered in case of changes of Clpc of more than 10% compared to the last Cltd”.

6. I found the Discussion a bit too long. In my view this could be shortened especially the part about the PiCCO algorithm, and also, practical implications should be practical, instead of repeating what was already explained before.

Answer: This is a very valuable comment. Therefore, we shortened the Discussion by several paragraphs related to
- other parameters derived from thermodilution (GEDVI; EVLWI)
- general knowledge on precision and accuracy (definitions of bias, precision, limits of
agreement etc.) and
- redundant parts of the discussion.

Ad Reviewer #2:

1. The results are mainly based on the correlation between Clpc-Cltd and various factors. How did the authors choose such factors?

This is a very valuable question. The study was aimed to define parameters associated to substantial inaccuracy and/or imprecision of Clpc compared to immediately subsequent Cltd. Indeed, correlation of potential variables ("predictors") to Clpc-Cltd were first steps of analyses which were completed by partial correlations, ROC-analyses, Bland-Altman analyses and Generalized Linear Mixed Models.

With regard to practical implications the study also tried to investigate the potential of automated decision support. Therefore, - in addition to time to last calibration and Cltd - the factors investigated were chosen among parameters continuously provided by pulse contour analysis. To clarify this, we included the following sentence in the Materials and Methods section: “These parameters included “time to last calibration” as well as factors continuously provided by pulse contour analysis and their changes compared to baseline”.

2. It is obvious Clpc-Cltd is correlated with Clpc, and with Clpc-Cltd (baseline), since Clpc is one of the two determinants of the Clpc-Cltd values. In other words, correlation may more result from mathematics than from hemodynamic factors.

This is a very interesting comment. Indeed, at first glance including Clpc in “Cltd-Clpc” and Clpc-Cltd(baseline) could suggest a correlation of “Cltd-Clpc” and Clpc-Cltd(baseline)”. However, this is not substantiated by mathematics.
To give an example: If A is a person´s birth-weight, B is the person´s weight at 79 years estimated by a third person and C is the true weight at the age of 80 years, the accuracy of the estimation (i.e. B-C) does not necessarily depend on the difference between the estimation and the birth weight (B-A) or on the estimate itself (B).

Another example: A is a person´s bank account balance is 1000€ at the age of 10 years. Another person estimates the account balance at the age of 79 years (B). The difference B-C between the estimate B and the true value at 80 years (C) does not necessarily depend on the estimate itself (B) or its difference to A (B-A).

To give a third, geographic example: If the distance from Munich to Budapest is 1000km (A) and the distance from Paris to Budapest is 2000km (C), being 900km away from Munich (B-A) does not necessarily mean to be 100km (C-B) away from Paris.

We hope that these examples illustrate that the accuracy of an estimation (CIpc) compared to the exact follow-up parameter CItd does not necessarily depend on the height of the estimate CIpc or its difference to a baseline value (CItd(baseline)).

Nevertheless, we are grateful for the Reviewer´s comment, and mathematic coupling has to be discussed. Mathematical coupling has to be considered when correlation of two variables A and B is improved by multiplication, division, subtraction or addition of both A and B by a third variable C (Walsh et al. [1]). The rational to check our findings for coupling could be that the dependent variable “CItd-CIpc” as well as the independent variable “CIpc-CItd(baseline)” both include CIpc. In case of mathematical coupling, the subtraction of CIpc of both CItd and CItd(baseline) would result in better correlation of CItd-CIpc and CItd(baseline)-CIpc compared to the correlation of CItd and CItd(baseline).
First, it is obvious that Clpc was subtracted from Cltd, but added to –Cltd(baseline). This could be overcome by changing the algebraic sign for one of the two correlated variables and –“in compensation” - also for the coefficient of correlation:

(I) The correlation of “Cltd-Clpc” and Clpc-Cltd(baseline) results in a coefficient of correlation $r$.

(II) Correlation of -1* (Cltd-Clpc) and Clpc-Cltd(baseline) results in a coefficient of correlation “$-r$”. 

Transcription of (II) results in

(III) Correlation of (–Cltd+Clpc) and (-Cltd(baseline)+Clpc) with a coefficient of correlation “$-r$”.

In (III) both “-Cltd” and “-Cltd(baseline)” are added by Clpc which could result in mathematical coupling. In case of mathematical coupling the amount of the coefficient of correlation should be higher for (–Cltd+Clpc) vs (-Cltd(baseline)+Clpc) compared to Cltd vs Cltd(baseline).

However, the hypothesis of mathematical coupling is rebutted by the data of our study: While correlation of Cltd vs. Cltd(baseline) in the merged data results in a coefficient of correlation of $r=0.910$ (p<0.001; partial correlation), correlation of Cltd-Clpc vs. Cpc-Cltd(baseline) results in a coefficient of correlation of $r=-0.606$ (p<0.001; partial correlation)). In other words, there was a strong correlation of Cltd and Cltd(baseline) before the addition of Clpc to these variables, while addition of Clpc markedly weakened the correlation. However, this is the opposite of mathematical coupling: As stated by Farmery [2], mathematical coupling “is described as a mechanism whereby correlation between variables is apparent in circumstances where none exists, and so cannot be supported by models in which correlation does exist”. This argumentation obviously applies for our data with a strong correlation of Cltd and Cltd(baseline).
3. The bias between CIpc and CItd may be associated with changes in arterial compliance. Although it is not possible to assess compliance at the bedside, one could perhaps approach it by calculating the ratio of stroke volume variation over pulse pressure variation. What would provide such an analysis?

This is a very valuable comment. Since parts of the measurements were performed with the PiCCO-plus-device, PPV was not available for all measurements. Analysis of the 260 datasets with both changes in SVV and changes in PPV available did not demonstrate an association of the bias CItd-CIpc with changes in the ratio of stroke volume variation over pulse pressure variation (r=-0.044; p=0.476). Furthermore, there was neither an association of baseline (r=0.043; p=0.434) nor of follow-up values of SVV/PPV (r=-0.006; p=0.927) to the bias CItd-CIpc.

Specific comments

1. Materials and methods. If I am right, each patient participated to the study at 8 instances: One time when the second TPTD was performed one hour after the baseline TPTD, another time when it was performed 2 hours avec baseline TPTD etc. If so, it is hard to understand in the present version. Please clarify.

This is a very good suggestion really improving the clarity of our paper. We now clarified in the Methods section that a total of 6 triplicate TPTDs at baseline and after intervals of 1h, 2h, 4h, 6h and 8h after the previous TPTD were recorded within 21 hours. Since follow-up TPTDs re-calibrated CItd, these measurements also provided the baseline CItd for the next interval.

2. The number of words of the abstract exceeds the number recommended by the Journal.
The word count of the Abstract without keywords is exactly 350 words which is in accordance with guidelines of both Critical Care and BMC Anesthesiology. Please let us know, if this exceeds the acceptable word count.

3. Please change the unit of measure of CI for “L/min/m²”.
This is a valuable suggestion. We changed the unit of any CI in the manuscript including Figures and Tables to L/min/m².

4. How was the sample size calculated?
There was no mathematical sample size calculation. The initial protocol was based on the inclusion of at least 50 datasets of at least 25 patients. After finding an unexpectedly high predictive capability of changes in CIpc compared to baseline CItd regarding the bias CItd-CIpc and with regard to the absence of an association of time after the last thermodilution to the bias we decided to validate our data in a second independent dataset with a similar size.

5. Results. “Mean bias values were -0.06061±0.60276 and 0.00261±0.60961 L/min*m², respectively.” Please reduce the number of decimals.
This is a very good suggestion. We now reduced the number of decimals to a maximum of three.

6. “Fig. 1 and Table 3 ... To avoid repeated measurements and different...” The second sentence is more related to methods than to results.
This is a very valuable comment. We now removed this sentence—which was already included in the Methods section—from the results section.

7. “Obviously, in both collective”, “bias-values also demonstrate”, “this is ‘strong evidence’ that”... Such terms interpreting the results should be removed from the Results paragraph and rather belong to the discussion.
This is a valuable suggestion. We removed “strong evidence”, “obviously” and “also”.

8. The discussion is too long and could be easily shortened.
This is a very valuable comment (see Reviewer 1; 6.). Therefore, we shortened the Discussion by removing several paragraphs concerning other parameters derived from thermodilution (GEDVI; EVLWI), general knowledge on precision and accuracy (definitions of bias, precision, limits of agreement etc.) and redundant parts of the discussion.

9. The discussion lacks hypotheses explaining why the present results are in contradiction with the previous ones.

In the Discussion section we clearly stated that “Driven by the practical need to define "when" to re-calibrate, time-dependency of Clpc-accuracy is an obvious hypothesis”. Furthermore, clinical practice and recommendations of the manufacturer are clearly based on the assumption that re-calibration should be performed after a maximum of 8 hours. Therefore, our findings and suggestions to associate re-calibration to changes in Clpc compared to the last Cltd are contradictory to the present practice.

10. Page 13: “In our merged data, percentage-error... None of these analyses provided evidence for time-dependency of the agreement of Clpc and Cltd.”

This is unclear to time. Please reformulate.

In this paragraph we summarized our findings and tried to emphasize that despite numerous statistical approaches including uni- and multivariate analyses in two different collectives as well as in the merged data there was no hint for time-dependency of the agreement of Clpc and Cltd.

Nevertheless, we reformulated and stated in the limitations section: “Although our data suggest that the (in)accuracy of Clpc is predominantly associated to changes in Clpc compared to baseline Cltd, we cannot definitely rule out a certain impact of time to last calibration due to the limited number of patients”.

We hope that this softening of our conclusion is acceptable for Reviewer 2.
Additional material submitted by the reviewers:

"The question of defining some factors that may be associated with recalibration is interesting but inadequately addressed -Some of the interventions should not be included in the analysis (prone/supine, cardiopulmonary resuscitation, tracheostomy?). In many instances, physicians/nurses would consider recalibration in these circumstances whatever the changes in CIpc. You should concentrate on more usual hemodynamic interventions (fluids, vasoactive agents). This means that the analysis was influenced by 40% of non hemodynamic interventions, and this may have a major impact on the different cut-offs.

-Unfortunately, focusing only on hemodynamic interventions (fluids, vasoactive agents) would decrease the number of datasets to only 71, which may be insufficient to drive strong conclusions. You should therefore increase the sampling size, focusing on relevant interventions.

Also computation or aortic elastance (DPP/SVV) should be implemented."

We thank Reviewer 2 for these additional comments and for the valuable suggestions for future approaches.

Indeed, we tried to give extensive information about the collective investigated in Table 1. This was in part aimed to support that this study was not exclusively observational. Furthermore, we tried to emphasize that these patients were critically ill due to a variety of etiologies.

We clearly admit that a larger number of patients investigated could have further supported our findings. Nevertheless, we feel that using a second validation collective as well as two different ICUs and including a total of 738 measurements (including the baseline measurement of the 1st interval) in 77 patients might
substantiate our findings and might improve insight in this topic. E.g. to the best of our knowledge, the study of Hamzaoui and co-workers [3] is considered as the current landmark study regarding re-calibration of Cltd with the largest collective investigated. This retrospective single-center study analyzed 400 measurements in 59 patients. Catecholamines were used in 29/59 patients (49%) which is the same percentage as in our study.

Furthermore, we do not feel that exclusion of patients who underwent interventions such as tracheotomy or prone positioning would strengthen our study. In our opinion a clinical study should be "as close to clinical reality as possible".

Finally, a “calibration alarm” suggesting re-calibration rather after substantial changes in Clpc than after a fixed time does not preclude re-calibration for other reasons (e.g. intermittent measurement of GEDVI and EVLWI).

We hope that you find our revised manuscript worth to be published in *BMC Anesthesiology*.

Sincerely

Wolfgang Huber, MD
References:

