Author's response to reviews

Title: Comparison of a new transcutaneous bilirubinometer (Bilimed(R)) with serum bilirubin measurements in preterm and full-term infants

Authors:

Tanja Karen (tanjakaren@yahoo.de)
Hans Ulrich Bucher (hub@usz.ch)
Jean-Claude Fauchère (jean-claude.fauchere@usz.ch)

Version: 2 Date: 31 August 2009

Author's response to reviews: see over
August 27th, 2009

Mrs
Natalie Pafitis MSc
Senior Scientific Editor
BMC-series Journals
BioMed Central

Londonm, WC 1X 8 HL

Ref.: MS 9172263062801076

Comparison of a new transcutaneous bilirubinometer (Bilimed) with serum bilirubin measurement in preterm and full-term neonates

Dear Mrs Pafitis, Dear Mr Hewit,

Enclosed please find our revised manuscript for publication in BMC Pediatrics. We have addressed each of the reviewer’s comments and hope that our manuscript is now acceptable for publication. We have enclosed our response to the Reviewer’s comments, keyed to the Reviewer’s questions. All changes as requested by the reviewers were made.

All authors have read and approve of the revised manuscript. Please contact me if you require further information.

Sincerely

Fauchère Jean-Claude
Reviewers' comments:
Reviewer 1

Reviewer's report:
General remarks
We have shortened the discussion as requested by deleting the extensive discussion of the other devices. The results are ‘negative’ as stated by the reviewer, but not so in a statistical sense. Our data are negative in that they show that this device does not perform as accurately as requested by clinicians taking care of newborn babies. In this aspect, some of our data are in line with other more recent studies on the performance of transcutaneous bilirubin measurements, but our data provide valuable information to the clinician regarding measurements with Bilimed® in non-Caucasians and in a larger gestational age scale.

1. Presentation of decision for the number of infants studied
The aim of the study was to enrol 100 term infants (maternity ward) and 50 preterm infants (maternity ward, intermediate care and NICU). As explained in the section ‘discussion’, the reading of the TcB was often very difficult, if not impossible, in the very premature infants. Although not intended from the beginning of the study, we subsequently sub-divided the group of preterm infants as there were clearly two different populations, both in terms of clinical presentation with the more immature sicker, and thereby also in terms of risk with regard to bilirubin encephalopathy. Based on their gestational age and for statistical analysis, the preterm infants were therefore allocated either to group 2 (34 0/7 to 36 6/7 wks GA) or to group 3 (28 0/7 to 33 6/7 wks GA). The technical difficulties encountered in getting acceptable TcB readings in a feasible time frame without disturbing too much the sicker and smaller infants led to the decision to stop enrolment of the most premature infants. This explains the fact that at the end of the study, there were more premature infants allocated to group 2 than 3.

2. Description of the method of serum bilirubin measurements and their variability.
We have clarified this point in the section ‘methods’.

3. Since three readings were taken from one site, it would be of interest to present the variability among these reading, e.g. the coefficient of variation
We have performed the calculations as requested, and we have included this information in the section ‘results’.

4. The data is presented with r values but no mention whether significance is present.
We have added the level of significance as requested in the section ‘results’.

5. Assumptions of the other devices improved performance comes not from head-to-head testing but from the literature. It’s possible that in their hands, in their institution, the other devices may have performed differently from that reported.
In this study, we did not perform a head-to-head comparison with other transcutaneous bilirubin measurement devices. But the authors have published research done with the BiliCheck and Minolta (references 8 and 10).

6. The rationale for using the chest instead of the forehead is given, but it would have been more convincing had there been a preliminary investigation comparing these sites.

We fully agree with the reviewer in that it would have been interesting to perform a preliminary study on the best sampling site for the BiliMed. We have based our choice on available data from the literature, clearly favouring the sternum over the forehead. Two further important arguments against using the forehead was that we chose to study more preterm infants who would possibly wear N-CPAP bonnets, or have very small sampling areas due to their body size. Another site for comparison could have been the back of the infant. We chose not to perform the measurements on this alternative site, and also not to compare it to the sternum, because of ease of access to the sternum and thereby not to add any discomfort to these newborn infants lying on their backs as it is the policy of our clinic (‘back to sleep’).

7. The report concludes with a statement of the advantages of this new device, clearly not appropriate since it doesn't seem to work very well.

We have changed the phrasing in the section ‘discussion’.

Reviewer 2

Reviewer's report:

1. You have shown the regression analysis in non-Caucasian babies. To be complete you MUST show the same analysis in the Caucasian subgroup, so that regression equation and device performance may be evaluable. I also advise you to form a panel of figures (i.e.: Fig.2A and B) in which put together the two graphs.

As requested by the reviewer, we have presented the regression analysis for both the Caucasian (Fig. 2 A), and for the non-Caucasian infants (Figure 2B).

2. I do not think worthless to perform a separate analysis on smallest babies, since they are just 13. Therefore why do not put together groups 2 and 3 and perform an unique regression? In that case you will have another panel with the fig.1A for full term infants and the fig.1B for preterm.

We fully agree with the reviewer’s proposal and we have therefore merged group 2 and 3 and shown the results in Figure 1B, with Fig. 1 A showing the results for the term infants.

3. When commenting your equation please state that the constant is far different from zero in preterm babies and in non-Caucasian. Will it be greater in non-Caucasian compared than in white babies? Please answer (see point 1
above). We have added this information in the section ‘discussion’.

4. You forgot to shown the Pearson’s correlation coefficient. This is the easiest tool to study the agreement of the two devices and is strongly related to the R2 value. PEARSON MUST BE REPORTED AND R2 VALUE MUST BE COMMENTED. In fact it is very very low in some babies (0.17 in preterm babies, that means that just 17% of serum bilirubin could be correctly predicted by the device), but even quite low in white full term babies (only 0.52, similar to my study [0.45]). These values (together with the Pearson) are the meaningful method to show how unreliable is this device and definitively needs to be commented in results and discussion.

   We have, as requested, added the Pearson’s correlation coefficient in the section ‘results’, and discussed these data in the section ‘discussion’.

5. A point of main interest is the practical difficulties you have encountered when using the device on the smaller babies. I DO BELIEVE THIS IS HIGHLY IMPORTANT, for evaluating a device whose performance is already low and which is intended to be more practical than others. Please ADD in the RESULT THAT YOU STOPPED THE ENROLMENT OF SMALLER BABIES DUE TO THESE PRACTICAL PROBLEM. PLEASE ALSO EXPAND THE SECTION ADVANTAGES/DISADVANTAGES (in discussion) and better detail these problems. Can you give us an estimation of the time (in seconds/minutes) needed to perform the measurements in VLBW compared to the bigger babies? Why do not include a picture of the instruments in use in a VLBW baby so that we can see their relative sizes? This would be highly attractive.

   We fully agree on the importance of this information regarding the most fragile infants at greater risk of the deleterious effects of bilirubin. We have therefore expanded this in the section ‘discussion’. We have a picture of showing the BM in a more preterm infant in the incubator, but unfortunately not in a VLBW infant. Moreover, we are no longer in possession of the Bilimed.

6. Your assumption that BiliCheck, and in general 2nd generation transcutaneous bilirubinometers, are strongly affected by skin colour and PIGMENTS IS NOT TOTALLY TRUE. (Pag.6, last par). They have been developed to overcome the interference due to other pigments but more and above this we have some data that demonstrate this ability. For example you cited the sound paper of Billy Engle (Pediatrics 2002, Ref.5) in which BiliCheck was not so reliable in Hispanic babies. Nevertheless, Billy recently published a nice paper in which Minolta JM-103 performed very well in this population (Am J Perinatol 2009). Moreover its first paper has been extensively criticised because of some enrolment criteria details which comment is beyond the aim of this review. You should read these comments in Schumacher RE, Pediatrics 2002 and Bhutani, Pediatrics 2002. The main problem is that no device is perfect if you ask a punctual value of serum bilirubin, because skin bilirubin is not serum bilirubin: skin bili is really another clinical variable (thought related with the first). Therefore these device are useful as screening tool and not to have a punctual value: this validity has now been recognised even in Black African (and so highly pigmented babies...see
Slasha et al. Pediatrics 2004). In this particular population, in fact, BiliCheck performed very well. YOU MUST MODIFY YOUR COMMENT IN DISCUSSION ACCORDING TO THESE CONSIDERATION AND CITE SOME OF THE PAPERS THAT I SUGGESTED YOU.

We have changed our comments and also stressed the message regarding the value of a TcB value as a screening tool in the section ‘discussion’ and ‘conclusion’.

7. THE SENTENCE “Furthermore and in contrast to the BiliCheck, we did not see any differences of the transcutaneous bilirubinometer readings with the Bilimed® under different ambient light condition [8].” MUST BE DELETED. In fact you have done no comparison in different light conditions. You have not measured room light intensity or hours at the time of sampling and so you have no data to support the reliability of Bilimed® in different light exposure. On the contrary, a recent paper of mine, demonstrates that BiliCheck is not affected by light intensity (De Luca D, Turk J Pediatr 2008). This was a study specifically directed at this point in which we analyzed skin bilirubin measurement, according to the time of measurement. You could be interested in read it unless this is not mandatory for the study aim.

We fully agree with the reviewer. As we did not investigate the effect of ambient light conditions on TcB readings, we have dropped this paragraph.

DISCRETIONARY REVISION
The following points do not require mandatory changes in your paper but would be nice to be considered and would much improve your article.

1. I understand that you use bili determination at your central hospital Lab. Sometimes this could be more expensive and surely more time-consuming comparing to a bed side determination with direct spectrophotometry (performed directly in the nursery). You should specify that the assay in the central Lab is your usual internal policy for assaying bili and that your system is enough fast and accurate to be practical and reliable (maybe your Lab technician could say you some reliability value of the method; i.e.: CV).

We have addressed the points raised by the reviewer in the sections ‘methods’ and ‘discussion’.

2. The paper is lacking of the details about statistical analysis. Please add that you calculated the Pearson coefficient and please add what software did you use and what was the significativity threshold accepted (alfa-error, 0.05 I bet).

We have added this statistical information in the section ‘methods’.

3. Finally, after having corrected graph and statistics as suggested, why do not you launch a multiple linear regression analysing the performance of Bilicheck for predicting serum bilirubin and having as covariate the gestational age (term vs preterm) and the ethnicity (Caucasian vs non-Caucasian). Your population is enough wide to consider these two confounding factors and this would be a more elegant and advanced way for analyzing the influence of these factors. This would complete your analysis and assumptions.
We did a multiple linear regression analysis showing that birth weight is a confounding factor but not skin colour, nor gestational age (section result and discussion).

Reviewer 3

Reviewer's report:
*The paper of Karen et al. compares a new transcutaneous bilirubinometer with serum bilirubin determination in neonates of various gestational age and different skin colours. The new device is working with a system of 10 LEDs and is based on the analysis of five wavelengths. The employed methodology, regression's line between transcutaneous and serum measurements, and Bland-Altman plots of all comparisons are adequate to analyze the results. The correlation's coefficients between Bilimed# and serum bilirubin concentrations were unsatisfactory as well as the results obtained by the method of Bland and Altman. The paper is clearly written, and deserves publication.*

We thank the reviewer for his appreciation of our research work.