Author's response to reviews

Title: Improving Accuracy in Assessing Pulmonary Edema on Bedside Chest Radiographs Using a Standardized Scoring Approach

Authors:

Matthias Hammon (matthias.hammon@uk-erlangen.de)
Peter Dankerl (peter.dankerl@uk-erlangen.de)
Heinz L Voit-Höhne (heinz.voit-hoehne@klinikum-ansbach.de)
Ferdinand J Kammerer (ferdinand.kammerer@uk-erlangen.de)
Michael Uder (michael.uder@uk-erlangen.de)
Rolf Janka (rolf.janka@uk-erlangen.de)

Version: 6 Date: 3 January 2014

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

we are very grateful for your thoughtful and helpful comments to our manuscript. We attempted to revise the manuscript in accordance with your advice. Here below is one-by-one response to your comments.

Editor's comment.

1) EVLW-values provide reliable evidence of a pulmonary edema; a score of $\leq 8$ mL/kg indicates no pulmonary edema, while a score of $\geq 15$ mL/kg indicates a pulmonary edema: please, provide evidence that this is true in the individual patient, or the sensitivity and specificity of this method (see Anesthesiology. 2009 Nov;111(5):1065-74, and Anesthesiology. 2009 Nov;111(5):933-5). There is substantial discussion on the use of PiCCO, and the authors should be more balanced when making such statements. Indeed, there is significant effect of the distribution of perfusion on the measurements. Unless the authors have an independent method to determine the accuracy of their statement, this should be limited to the evidence provided by the literature. This should be accurately cited.

Thank you for your comment and the very interesting literature you provided. We thoroughly studied the literature and balanced our statements accordingly. We defined an EVLW score of $\leq 8$ mL/kg as no pulmonary edema because this is clinical practice in our ICU. There is a limited number of studies on this issue. An animal study in dogs shows an physiologic EVLW value from 3-7 ml/kg (Ann N Y Acad Sci. 1982;384:394-410 = Reference 28).

We rephrased and extended the corresponding paragraphs in the Background and the Methods section.
Invasive techniques such as the indicator dilution method which measures the extravascular lung water using a PiCCO® system (Pulse Contour Cardiac Output; Pulsion Medical Systems, Munich, Germany) seem to be reliable for evaluation of pulmonary edema in patients with indirect lung injury [2], but they are expensive and there is a risk of complications developing.

Although there is discussion on the use of PiCCO measurements, especially regarding the normal clinical range of EVLW as well as on the effect of the distribution of perfusion on EVLW values [17, 18], there is evidence that the EVLW-value is a sensitive marker of pulmonary edema [19 -21]. Roch et al. provided evidence in an animal model that EVLW measurements are useful and reliable for evaluation of pulmonary edema in indirect lung injury, but produce misleading values in direct lung injury [2]. As all patients included in this study show indirect lung injury, we assume that the determined EVLW values are reliable. We defined an EVLW score of ≤ 8 mL/kg as no pulmonary edema, and a score of ≥ 15 mL/kg as pulmonary edema [19, 22 - 28]. Patients presenting a borderline EVLW-value (9 – 14 mL/kg) and who, therefore, could not be reliably classified, were not included. A sequence of ten patients with an EVLW-value of ≤ 8 mL/kg and a series of ten patients with an EVLW-value of ≥ 15 mL/kg were included. In addition, patients were selected without prior knowledge of their clinical diagnosis.

2) Please, provide the confidence limits for the estimates entered in Table 2 (sensitivity, specificity. etc.)

Confidence limits were provided as demanded (Table 2).
Pending Issues on the answers to the Associate Editor’s Comments:

(c) Clarify how the authors selected the criteria or weighting of the criteria for diagnosing pulmonary edema: the answer provided by the authors to this question is insufficient. This is because it does not justify the specific variables and scores utilized.

1. Why did the authors choose those variables and scores?
2. Have they performed an analysis of colinearity to determine that those variables are independent?
3. How do they know, for instance, that enlarged hilar vessels are worth a maximum of 3 points while diffuse increased density is worth 15 points?
4. Where did the numbers 3 and 15 (as all the others) come from?
5. Was this, for instance, the results of previous multivariate analysis?

Although the bedside chest radiograph is the most frequently performed conventional X-ray you can hardly find literature on the assessment of pulmonary edema. These criteria and scores to assess pulmonary edema on bedside CXRs were validated previously [1, 13, 29, 30]:

The weighting of the different criteria depends on how distinctively EVLW is shown by each specific criterion. For example, enlarged hilar vessels (maximum of 3 points) may have multiple causes but they do not show EVLW. In contrast, extensive perihilar haze (maximum of 15 points) is caused by and actually represents EVLW.

A paragraph was added

[Customized scoring software, implemented as a web-based application, facilitates a standardized assessment of the presence of pulmonary edema. It is based on the systematic evaluation of characteristic radiological indications of pulmonary edema on bedside CXRs which was previously validated [1, 13, 29, 30].]

(d) provide inter-reader or intra-reader variability?: A consulted statistician thinks that calculating intrareader variability would not reasonable in this context.

Please, scientifically justify this statement. I do not see why an intrareader variability is not justified. If you mean that there should not be a different variability in comparison to the average intrareader variability of radiologists, please, provide references for that variability. If available, please, provide evidence supporting that the variability in the reading of edema is equivalent to that in other cases.

We provided interreader variability to show that a standardized assessment of pulmonary edema on bedside chest radiographs leads to a higher degree of agreement among radiologists. Free-marginal multirater kappa values were calculated for the standard as well as the standardized score-based approach. Values were included in Table 2 and mentioned in the Results section.
Intrareader reliability is the degree of agreement among multiple repetitions of a diagnostic test performed by a single rater. As every radiologist rated each radiograph just once with the standard as well as with the score-based approach, we are unfortunately not able to provide intrareader variability. In order to calculate intrareader variability every reader would have to assess the radiographs multiple times under steady (following the same rules) conditions.

(e) it is difficult to understand why this is not feasible. If the authors have the radiologists’ reading, they should be able to know which of the criteria best related to identification of edema. This is an important aspect of the analysis. Please, let me know if this question is not clear and I would be glad to explain to it.

You mention a very interesting point.

Table 1 shows weighting of the different criteria (maximum score of 3 - 15). A higher maximum score means that the corresponding criterion shows extravascular water more distinctively. As enlarged hilar vessels do not actually show extravascular water and because this sign may have multiple causes the maximum score is 3. In contrast, extensive perihilar haze or a diffuse increase in density is caused by and actually represents extravascular water and therefore got a maximum score of 15 points.

Due to the fact that we only assessed 10 bedside CXRs showing criteria of pulmonary edema we are not able to establish which imaging criteria best correlate with the EVLW categories. This issue should be addressed in ongoing research.

We added a paragraph to the Limitations section covering this point:

Due to the fact that we only assessed 10 bedside CXRs showing pulmonary edema we are not able to establish which imaging criteria best correlate with the EVLW categories. However, table 1 shows weighting of the different criteria based on findings of previously
published research [1, 13, 29, 30]. A higher maximum score means that the corresponding criterion shows extravascular lung water more distinctively. For example, the criterion “enlarged hilar vessels” got a maximum score of 3 points because it may have multiple causes and it does not show extravascular water indeed. In contrast, the criteria “extensive perihilar haze” and “diffuse increase in density” are caused by and actually represent extravascular water and therefore got a maximum score of 15 points.

Thank you again for your kind and thoughtful review. We hope that the revision addresses your concerns and that you will find the manuscript up to the standards of *BMC Anesthesiology*. We very much look forward to hearing from you.

Sincerely,

the authors