Author's response to reviews

Title: A latent class approach for sepsis diagnosis supports use of procalcitonin in the emergency room

Authors:

Fabián A Jaimes (fjaimesb@jhsph.edu)
Gisela De La Rosa (giseladlr@une.net.co)
Marta L Valencia (martaluzyvalencia@gmail.com)
Clara M Arango (clara.arango@gmail.com)
Carlos I Gomez (cigr@une.net.co)
Alex Garcia (deorumalex@yahoo.com.ar)
Sigifredo Ospina (soox@elhospital.org.co)
Susana C Osorno (susana_upegui@yahoo.com)
Adriana I Henao (adrisahenao@yahoo.com)

Version: 3 Date: 16 May 2013

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

We are resubmitting the manuscript entitled: “A latent class approach for sepsis diagnosis supports use of procalcitonin in the emergency room” considering the comments and corrections requested by the reviewers. Below, we are providing a point-by-point response to the concerns. In addition, specific changes are underlined in the manuscript

Reviewer 1

Minor Essential Revisions

The agreement between experts is low. A rate of 0.65 for sepsis-no sepsis and 0.73 for infection with and without sepsis is strongly undermining the validity of data. The authors should provide a good reasoning for the above and acknowledge it as a significant limitation.

Thank you very much for your comments.

Certainly, the agreement among experts for clinical diagnosis of sepsis and/or infection is low. But this finding, indeed, underlines the problems related to the usual process of diagnosis for this condition. Consequently, such a clinical disagreement supports the search for alternative methods in the study and analysis of sepsis. We realize in the first paragraph of the discussion this issue: “Indeed, the kappa-statistic measure for multi-rater agreement between experts for this definition was 0.65 for sepsis-no sepsis and 0.73 for infection with and without sepsis, which underlines the limitations for clinical diagnosis in this condition”.

In addition, we add the following paragraph below the issue of limitations. “As we mentioned before, the clinical diagnostic “gold standard” utilized here performed poorly, as the concordance between experts was 0.65 for sepsis-no sepsis and 0.73 for infection with and without sepsis. This weakness, indeed, underlines the limitations for clinical diagnosis in this condition.”

The ‘cluster’ of 187 patients should be described in detail. Are these ‘cluster’ patients, patients with ‘severe sepsis’ or/and ‘septic’ shock? Have patients with severe sepsis and septic shock according to the traditional criteria been included in the LCA gold standard cluster (the severity of the cluster patients seems higher)? If not, could the authors provide the ‘identity’ of these cluster patients, as well as the ‘phenotype’ of the missed severe sepsis pts.

This “cluster” of 187 patients is described in detail in table 3. This population it is more severely ill and they also have more frequently positive blood cultures. Furthermore, we add in the results the following information: “According to standard definitions, 70% (n = 131) of these patients had severe sepsis without circulatory failure and 5% (n = 9) had septic shock (Table 3)”.

LCA approach carries methodological novelty in the field of infectious diseases. However the important question is whether this approach carries any advantage compared to the traditional one. Sepsis is not an illness, but a syndrome defined mainly on clinical criteria and the misdiagnosis of sepsis is associated with an extremely adverse outcome. The authors should discuss the above issues and underscore the issue of safety.

You are absolutely right: “Sepsis is not an illness, but a syndrome defined mainly on clinical criteria and the misdiagnosis of sepsis is associated with an extremely adverse outcome”. However, we are not proposing a new methodological approach for sepsis diagnosis. We are identifying, provided the clinician requests a procalcitonin test and the results are rapidly available, a new cut-off point to detect more severely ill patients. This goal, as we showed in the other results and analyses in the paper, was not achieved by the conventional “clinical only” approach. In this way, LCA is just an instrument to show that we can improve the process of sepsis diagnosis in the emergency room if consider a higher value for procalcitonin. We add in the discussion the next paragraph: “Needless to say, sepsis is not an illness but a syndrome suspected mainly on clinical criteria, and the misdiagnosis of sepsis is associated with an extremely adverse outcome. Consequently, we are not proposing a new methodological approach for sepsis diagnosis. Instead, we are identifying a new cut-off point for procalcitonin to be able to detect more severely ill patients. This goal was not achieved by the conventional clinical gold standard with expert consensus and, in this way, LCA is just an instrument to show that we can improve the process of sepsis diagnosis in the emergency room”. However, at the end of the discussion and below the item of limitations, we underscore the following: “On the other hand, LCA also has its limitations as "gold" standard. Under this approach, sepsis is not formally defined but rather is a mathematically defined entity that does not necessarily correspond with a clinically relevant status. Additionally, LCA modeling requires sophisticated analytic techniques and software, and the full model or the hypothetical “true” state of disease cannot be fully tested with the observed data”.

Given that 89% of pts had suspected infection as admission diagnosis, could the addition of clinical indices such as fever offer a better diagnostic performance following the LCA approach?

Among these 683 (89%) patients with suspected infection just 25% (n = 171) had fever at admission, which means that clinicians made their judgments most of the time considering additional factors other than fever. On the other hand, among these 171 patients with suspected infection plus fever just 75% (n = 128) was finally considered as sepsis by the expert committee. This percentage seems not particularly different of the 66% (n = 505) classified with sepsis among the total study population. Therefore, it seems that addition of clinical findings as fever provides no better diagnostic performance, either with or without LCA.
Reviewer 2

Major compulsory revisions

1) there is a putative bias in the population studied, as indicated by the very high percentage (70%) of patients having received corticosteroid and/or chemotherapy in the 3 months before inclusion, letting suggest that most of them had cancer, a condition known to be associated with chronic elevated serum levels of both CRP and DD. This may have underestimated the diagnostic value of these biomarkers. May the authors comment this ?.

Thank you very much for your comments.

We are very sorry about the confusion, as the number of patients with use of corticosteroids or chemotherapy during the past 3 months was 70, which corresponds to 9% (not 70%) of the study population.

2) The LCA allowed the identification of a more severely ill/bacteriologically documented subgroup of septic patients, therefore questioning the usefulness of any biomarker for these patients who have probably a more obvious sepsis presentation. Was the agreement between the experts significantly better for the 187 patients than that for the whole cohort ? If yes, this could argue against the need for a biomarker to identify them, as clinical presentation may be sufficiently suggestive of sepsis. How many of these 187 patients had severe sepsis/septic shock criteria?

This is a very interesting and challenging question. We reanalyzed the agreement between experts in the LCA-cluster, and the results were very particular:

```
.kap clas_eval1 clas_eval2 clas_eval3
There are between 0 and 3 (median = 3.00) raters per subject:

Two-outcomes, multiple raters:
   Kappa       Z       Prob>Z
-----------------------------
  0.7334 29.81  0.0000

.kap clas_eval1 clas_eval2 clas_eval3 if clcluster==1
There are between 0 and 3 (median = 3.00) raters per subject:

Two-outcomes, multiple raters:
   Kappa       Z       Prob>Z
-----------------------------
  0.7526 15.01  0.0000
```
There are between 0 and 3 (median = 3.00) raters per subject:

Two-outcomes, multiple raters:

<table>
<thead>
<tr>
<th>Kappa</th>
<th>Z</th>
<th>Prob&gt;Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0.6386</strong></td>
<td>31.32</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

There are between 2 and 3 (median = 3.00) raters per subject:

Two-outcomes, multiple raters:

<table>
<thead>
<tr>
<th>Kappa</th>
<th>Z</th>
<th>Prob&gt;Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0.4792</strong></td>
<td>11.51</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

As you can see, they were overly similar for infection with and without sepsis (0.73 vs. 0.75) but were even worse for sepsis-no sepsis (0.65 vs. 0.48). For better clinical characterization, finally, we add in the results the following information: “According to standard definitions, 70% (n = 131) of these patients had severe sepsis without circulatory failure and 5% (n = 9) had septic shock (Table 3)”.

Minor essential revisions

The main result is that PCT is useful to identify a subgroup of more severely ill septic patients attending the ED. This has already been published and must be more extensively discussed in discussion’s section. (Serum procalcitonin measurement as diagnostic and prognostic marker in febrile adult patients presenting to the emergency department. Hausfater P et al. Crit Care. 2007;11(3):R60).

That is correct. We incorporate in the discussion the following text: “Our main result, consequently, is that PCT is useful to identify a subgroup of more severely ill septic patients attending the ER. Such a finding was previously reported by Hausfater P et al., (Hausfater P, Juillien G, Madonna-Py B, Haroche J, Bernard M, Riou B. Serum procalcitonin measurement as diagnostic and prognostic marker in febrile adult patients presenting to the emergency department. Crit Care. 2007;11(3):R60) whom studied 243 patients with body temperature of 38.5°C or greater attended in the adult emergency department of an academic tertiary care hospital. They found, using standard statistical methods, that PCT is an independent variable that can predict whether a febrile episode has a bacterial origin, and that at a threshold of 2 µg/l it is independently associated with critical illness. The coincidence with our findings is remarkable, despite the fact that their study population was extremely different: all the patients consulted by a febrile episode, 29% of them were immunocompromised, and only 81% were hospitalized in that consultation.”
Reviewer 3

Major issues: 1) My strongest personal issue is the mixture of both story lines. Firstly, the authors used LCA to overcome the lack of a gold standard and to identify patients with a “higher degree of sepsis or a more severe progression of disease”. This test based among others of PCT, CRP and DD levels. Afterwards, they defined in this “more severe ill” patient group new cut of values for PCT. As far as I understood the method, LCA included higher levels of PCT, CRP and DD to define the group of “more severe ill patients”. Therefore, after this “pre-testing” higher sensitivities and specificity are likely. Further, it is not clear how these new cut-offs are helpful for physicians since they have to use the new cut-offs in the overall not pre-tested patients cohort. Therefore, I would recommend to focusing on LCA as new diagnostic test for sepsis and to avoid a overstressing of the ROC-PCT results. Higher PCT results are useful in terms of confirmation of your LCA.

Thank you very much for your comments. We are very sorry about the confusion, and we try to explain the issue better in both, the methods and discussion sections. LCA, strictly speaking, is not a new diagnostic test or even a diagnostic method for sepsis. This is just a mathematical model, which identifies a subtype or a cluster of observations according to certain defined characteristics or variables that are common to those observations. In this case, we know that different expressions of inflammation and coagulation are common responses in the process of infection. We provide these observed variables (D-dimer, Procalcitonin and C-reactive protein) from all the study population to the model and it is able to uncover the hidden group, i.e. the latent variable, to which the patients belong. Once this “latent class” of sepsis patients is detected, and we assume that they represent those confirmed by the “new and best” but just mathematical gold standard, we found that PCT provide the strongest discriminatory power among those proposed biomarkers (Table 5), but with a higher cut-off point of 2 ng/ml. Neither this new cut-off point nor the subgroup of severe sepsis patients were identified with the classical Bayesian analysis using three independent experts as gold standard. Therefore, PCT > 2 ng/ml is not only a confirmation of the LCA results. This is a clinical tool that suggests to the physician that a patient with suspected infection, any not pre-tested patient but with suspicion of infection in an ER, deserves special attention.

2) Without any personal association to the paper of Wacker C et al. in Lancet Infect Dis. I recommend to include and discuss the results of this meta analyses even though this paper did not focus on the ER setting.

We add the following paragraph in the discussion: “In a recent meta-analysis (Wacker C, Prkno A, Brunghorst FM, Schlattmann P. Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis. Lancet Infect Dis. 2013 May;13(5):426-35), Wacker C. et al. analyzed 30 reports, although only two from ER, accounting for 3244 patients. Bivariate analysis yielded a mean sensitivity of 0.77 (95% CI = 0.72–0.81) and specificity of 0.79 (95% CI = 0.74–0.84) and the area under the receiver operating characteristic curve was 0.85 (95% CI = 0.81–0.88). The median cut-off for PCT of the studies included was 1.1 ng/mL (IQR = 0.5–2.0) and the absence of a threshold effect suggests that a cut-off of between 1.0 and 2.0 ng/mL, close to our findings, is helpful for discrimination of patients
with sepsis from other inflammatory conditions. However, the studies had substantial heterogeneity ($I^2 = 96\%, 95\% \text{ CI} = 94–99$) and none of the subgroups investigated, like population, admission category, assay used, severity of disease, and description and masking of the reference standard, could account for that heterogeneity. They concluded that the test may be helpful for diagnosis of sepsis in critically ill patients, but it must be interpreted in context with information from careful medical history, physical examination and microbiological assessment.

3) I would recommend to characterize patients after re-classification by LCA in a overview using a Venn-diagram to compare the results to the clinical definition of sepsis.

This could be useful if we had seen any overlap between the cluster of sepsis and the clinical classification provided by the experts. However, these 187 patients were all among the group of 505 patients called sepsis by the consensus committee.

Minor issues:

1) I would recommend not to state LCA as one of the two gold standards for sepsis diagnosis as done in the abstract.

We rewrote this section in the abstract: “cross-sectional study to determine the diagnostic accuracy of three biological markers against the gold standard of clinical definition of sepsis provided by an expert committee, and also against the likelihood of sepsis according to LCA”.

2) Table 1. Please include p values. How do you explain the high rates of suspected infections in patients classified as non infected? The 28-day mortality rates should be presented more clearly (add %).

We corrected Table 1 with the percentages and the p values. Most of the patients (89%) were admitted to the hospital by the attending physicians with suspected infection, thus the distribution of these “misdiagnosis” should not be different compared to the whole cohort. As a matter of fact, the percentage of undetermined source of infection was 17% in these patients, compared to 3% in patients infected without sepsis and 13% in patients with sepsis.

3) Table 2. can be excluded.

Ok. We excluded it.

4) Table 4. Please include headlines in this table to clarify the meaning of cluster 1 and cluster 2. Please also include p values. I would recommend to include the levels of PCT, CRP and DD in this table.

Table 4 (now Table 3) was corrected.
# The authors refer to the study as “Prospective…cross-sectional”. This statement needs to be revised to accurately reflect the study design

Thank you very much for your comments. We rewrote the design as “Prospective single center study on the diagnostic accuracy of a test”.

# While using LCA as an analytical approach may be unique, the findings are not particularly novel.

You are probably right. However, a novel analytical approach that confirms the potential diagnostic and prognostic properties of a biomarker for the complex puzzle of sepsis is always welcome.

# Some discussion of the validity (or lack there-of) of the individual biomarkers selected and why these three markers specifically were chosen needs to be included. With the exception of procalcitonin, CRP and Ddimer have not been individually validated in bacterial infection and/or sepsis, and so it is not surprising they would perform poorly even when combined. While the authors provide some discussion of the prior literature on PCT, no explanation is offered about the lack of agreement between their identified cut-points and those previously published.

In the introduction we provided the references (15 to 19) and the ideas supporting the use of these biomarkers as potential candidates for sepsis diagnosis: “On the other hand, sepsis is associated with the simultaneous activation of the inflammatory and coagulation cascades, and most of their components are markers or mediators in the host response [12, 13]. From this close interplay between inflammation and coagulation, which is a recognized way toward organ dysfunction and mortality [14], emerges the rationale to characterize the host response to infection. Three potential biomarkers have shown regular presence in systemic infections: C-reactive protein (CRP), procalcitonin (PCT), and D-dimer (DD); the latter as an unspecific signal of coagulation activation [15-19].”

Regarding the cut-off point for PCT, we add the following paragraphs in the discussion: “Our main result, consequently, is that PCT is useful to identify a subgroup of more severely ill septic patients attending the ER. Such a finding was previously reported by Hausfater P et al., (Hausfater P, Juillien G, Madonna-Py B, Haroche J, Bernard M, Riou B. Serum procalcitonin measurement as diagnostic and prognostic marker in febrile adult patients presenting to the emergency department. Crit Care. 2007;11(3):R60) whom studied 243 patients with body temperature of 38.5°C or greater attended in the adult emergency department of an academic tertiary care hospital. They found, using standard statistical methods, that PCT is an independent variable that can predict whether a febrile episode has a bacterial origin, and that at a threshold of 2 µg/l it is independently associated with critical illness. The coincidence with our findings is remarkable, despite the fact that their study population was extremely different: all the patients consulted by a febrile episode, 29% of them were immunocompromised, and only 81% were hospitalized in that consultation.”
And before state our limitations: “In a recent meta-analysis (Wacker C, Prkno A, Brunkhorst FM, Schlattmann P. Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis. Lancet Infect Dis. 2013 May;13(5):426-35), Wacker C. et al. analyzed 30 reports, although only two from ER, accounting for 3244 patients. Bivariate analysis yielded a mean sensitivity of 0.77 (95% CI = 0.72–0.81) and specificity of 0.79 (95% CI = 0.74–0.84) and the area under the receiver operating characteristic curve was 0.85 (95% CI = 0.81–0.88). The median cut-off for PCT of the studies included was 1.1 ng/mL (IQR = 0.5–2.0) and the absence of a threshold effect suggests that a cut-off of between 1.0 and 2.0 ng/mL, close to our findings, is helpful for discrimination of patients with sepsis from other inflammatory conditions.”

The diagnostic “gold standard” utilized here, clinical expertise, performed poorly (65% agreement among cases). While the author’s acknowledge the lack of more powerful diagnostic tools in sepsis, this needs to be identified as a weakness.

We add the following paragraph below the issue of limitations. “As we mentioned before, the clinical diagnostic “gold standard” utilized here performed poorly, as the concordance between experts was 0.65 for sepsis-no sepsis and 0.73 for infection with and without sepsis. This weakness, indeed, underlines the limitations for clinical diagnosis in this condition.”

Please note that the complete corrected manuscript was reviewed and corrected by a native English speaker.

Sincerely yours,

Fabian Jaimes MD, MSc, PhD
Professor of Medicine, Clinical Epidemiology and Critical Care
School of Medicine
Universidad de Antioquia
Medellin, Colombia