PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multi-national, observational study
AUTHORS	Eckart, Andreas; Hauser, Stephanie; Kutz, Alexander; Haubitz, Sebastian; Hausfater, Pierre; Amin, Devendra; Amin, Adina; Huber, Andreas; Mueller, Beat; Schuetz, Philipp

VERSION 1 – REVIEW

REVIEWER	Jesus F Bermejo-Martin Group for Biomedical Research in Sepsis, BioSepsis Hospital Clínico Universitario de Valladolid. Spain I work in the field of biomarker discovery in sepsis and penumonia for a non profit organization.
REVIEW RETURNED	23-Jun-2018

GENERAL COMMENTS	This is a secondary analysis of data from a prospective, multinational, observational cohort study aimed to assess the prognostic performance of NEWS or NEWS + WBC, PCT, and MRproADM in the ER. The authors consider two outcomes, 30-day mortality and ICU admission. In both scenarios supplementing NEWS with the biomarkers chosen improves the performance of the former. These improvement seems to occur at expenses mostly of ProADM. This is a large study which opens an interesting avenue for improving the performance of clinical scores of severity such as NEWS. The work is well designed and the article is well written. This reviewer has nonetheless two observations that the authors may want to address: - Have the authors considered investigating the impact on their results of adding individual leukocyte subpopulations to NEWS? I understand that total leukocyte counts is the most frequently available data in retrospective analysis like this, but, if neutrophil or lymphocyte counts are available, I would suggest providing these additional analysis. - I suggest performing a subanalysis on the performance of the modified "NEWS + biomarkers" scores depending on the cause of admission (Infectious, cardiovascular, Pulmonary), as the authors did in Figure 1A for NEWS alone. Results for mortality prediction seems to heavily rely in the infectious disease group....The authors should discuss their results depending on these causes of admission.

	Minor comment: please do not use so many "moreover" in the intro section...

REVIEWER	Signe Søvik, Senior consultant anaesthesiologist, Associate professor Dept. of Anaesthesia and Intensive Care Akershus University hospital Norway and Institute of Clinical Medicine University of Oslo Norway
REVIEW RETURNED	01-Jul-2018

GENERAL COMMENTS	Thank you for allowing me to review this manuscript. It reports findings from a large, unselected, adult medical ED population included at three tertiary centres in Switzerland, France, and the USA, respectively. The study evaluates whether measurements of three inflammatory biomarkers (white blood cell count WBC, procalcitonin PCT, and mid-regional pro-Adrenomedullin MR-proADM) sampled upon patient arrival in the ED can improve the predictive power of NEWS scored upon patient arrival in the ED, outcomes being all-purpose 30-day mortality and ICU admission at any time during the hospital stay. The biomarkers had been prospectively sampled in a large study (TRIAGE), comprising 7132 patients. This study presents findings from a subsample of 1303 patients where sufficient data was available to retrospectively calculate a NEWS score. The study is well designed, succinctly presented, and the findings seem robust. Inherent limitations are mostly well addressed. Below are some comments and suggestions that I hope may contribute to further improve the manuscript. **Title** Inflammatory markers were sampled prospectively, but NEWS components were retrospectively collected, leading to a loss of almost 82% of the original patient population. I therefore strongly suggest removing the term “prospective” from the title. **Abstract** Page 3 Line 54-56: I suggest somehow rewriting the sentence “Combining the three inflammatory markers with NEWS improved risk stratification with regard to ICU admission compared to NEWS alone from AUC 0.70 vs. 0.65. “ e.g. “Compared to using NEWS alone, combining the three inflammatory markers with NEWS improved prediction of

	ICU admission, AUC increasing from 0.65 to 0.70." or "Combining the three inflammatory markers with NEWS improved prediction of ICU admission (AUC 0.70 vs. 0.65 when using NEWS alone)." Page 3 Line 50 and onwards See comments below regarding statistical tests assessing whether AUC values are significantly different **Introduction** Page 6 Line 26: Abbott, Ref 12, does not seem to have studied lactate. Is this an erroneous citation? **Methods** Page 10 Line 13: The applied NEWS categories (low ≤4, moderate 5–6 and high ≥7) are the ones recommended by the Royal College of Physicians. However, in the present study data on use of supplemental oxygen (worth 2 points) were unavailable. Thus many of the patients in reality would have received a 2 point higher NEWS score. It is likely that the fraction of patients that were under-rated by 2 points due to missing information on oxygen supplements would increase from the Low via the Moderate to the High NEWS group. This could represent at systematical bias where sicker patients (probably with poorer outcomes) were assigned too low NEWS values, possibly resulting in statistical inflation of the effect of the lower NEWS values. I do not think this limitation distorts the study's overall findings, i.e., obvious "dose-response" effects of NEWS. Also, the authors discuss this matter towards the end of the paper. However, I recommend underlining these facts and their potential effects a little stronger in Methods. Footnotes to figures and tables should mention that NEWS in this study was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2". In its present form, the text does not distinguish between actual NEWS values reported from other studies and the present study's modified NEWS calculations. Page 11 Line 47: I was unable to find mean and SD being used anywhere in the text. To be deleted? Page 11 Line 49: Replace "interquartile range (IQR)" with "quartiles" or "25th and 75th

	percentiles" throughout manuscript and tables. IQR denotes a single number describing the _distance between_ the 25th and the 75th percentile. Page 11 Line 49: What statistical tests were used to assess group differences referred with p values in Table 1? Please add information in Methods and as footnote to table. Page 11 Line 56: Was age represented as a linear covariate or binned as categories? Describe briefly how main diagnoses and comorbidities were coded and used in the statistical models. Page 12 Line 22: I suggest that you add a statistical evaluation of whether AUC values from the various models did in fact differ from each other. I believe STATA has routines for such calculations. **Results** Page 13 Line 8-11: I suggest adding "…. complete information for calculation of NEWS (excluding data on supplemental oxygen)…." What was the number of included patients from each of the three study sites, originally and in the present study? Page 13 Line 24: Does this AUC value (0.73) refer to the univariate model where NEWS alone predicts 30-day mortality, or the fully adjusted multivariate model where NEWS together with age, sex, main diagnosis, and comorbidities predict 30-day mortality (referred to in the previous sentence)? It might be useful to state in Methods (e.g., Page 12 Line 17) if presented AUC calculations systematically were calculated for models fully adjusted for age, sex, main diagnosis, and comorbidities, or systematically were calculated for models only including NEWS +/- inflammatory markers. The latter is how I understand the text. Page 17 Line 16 and Table 3: "… significantly improved the predictive value…" I suggest using a statistical test to compare AUC values. Many of the 95% CI's in Table 3 overlap, thus it is not obvious to the reader what models

	actually differed from each other. Please add information on statistical comparison of AUC values to Methods (e.g., Page 12 Line 20 and on) if such tests were performed, or if they are added to the study. If AUC values of the various models are not formally compared I suggest not using the term “significantly improved” since it implies a statistical evaluation. **Table 1** Page 14: Add a footnote on what statistical tests the p-values refer to. Add a footnote informing readers that NEWS was calculated without supplemental oxygen data. **Table 2** Page 16 Line 31: “CRP” should probably be exchanged with “PCT and Pro-ADM”? Add a footnote informing readers that NEWS was calculated without supplemental oxygen data. **Table A2** Page 34 Line 31: “CRP” should probably be exchanged with “PCT and Pro-ADM”? Add a footnote informing readers that NEWS was calculated without supplemental oxygen data. **Table A3** Page 35 Line 31: “CRP” should probably be exchanged with “PCT and Pro-ADM”? Add a footnote informing readers that NEWS was calculated without supplemental oxygen data. **Discussion** Page 21 Line 47: I suggest specifying in text, e.g., “In contrast, WBC results are available rapidly, and indeed PCT point-of-care tests that provide

	results within minutes are being developed". What would be the average time required to get results from an MR-proADM analysis, in a real-life ED setting? ED patients may be in a dynamic situation, and NEWS would be expected to change for the better in patients responding favourably to treatment. Clinicians would likely evaluate results from biomarker assays in the light of the patient's condition at the time the test results become available, not at the time the blood sample was drawn. Confer the findings of Abbott et al. (reference [12]) on NEWS values pre-hospitally and upon arrival in the hospital. Please discuss briefly. Page 22 Line 10: Although inflammatory markers were collected prospectively in the original patient population, the present study was an analysis of NEWS calculated retrospectively, in a population where it had not been ensured that all clinical measurements necessary for the calculation of NEWS was performed systematically. This resulted in a loss of almost 82% of the original patient population. Further, a modified NEWS still had to be used because data on supplemental oxygen was unavailable. Thus, I strongly suggest rewriting the term "secondary analysis of a prospective study" to "retrospective study" in the manuscript title and elsewhere. Page 22 Line 17: Please elaborate. How could batch measurements of blood samples address selection bias? Was not the problem that patients where clinicians had decided (more or less on purpose) not to record physiological data necessary for a NEWS calculation could not be included (=selected)? My apologies if I misunderstand your text. Page 22 Line 43 I do agree that multinational inclusion of patients is a strength of this study, and therefore recommend that the number of included patients from each site is stated in Results.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Jesus F Bermejo-Martin

Institution and Country: Group for Biomedical Research in Sepsis, BioSepsis, Hospital Clínico Universitario de Valladolid. Spain

General Considerations

This is a secondary analysis of data from a prospective, multinational, observational cohort study aimed to assess the prognostic performance of NEWS or NEWS + WBC, PCT, and MRproADM in the ER. The authors consider two outcomes, 30-day mortality and ICU admission. In both scenarios supplementing NEWS with the biomarkers chosen improves the performance of the former. These improvement seems to occur at expenses mostly of ProADM.

This is a large study which opens an interesting avenue for improving the performance of clinical scores of severity such as NEWS. The work is well designed and the article is well written. This reviewer has nonetheless two observations that the authors may want to address:

Specific points

1. Have the authors considered investigating the impact on their results of adding individual leukocyte subpopulations to NEWS? I understand that total leukocyte counts is the most frequently available data in retrospective analysis like this, but, if neutrophil or lymphocyte counts are available, I would suggest providing these additional analysis.

Reply: thank you for pointing this out. Unfortunately, measurement of leukocyte subpopulations was not part of the routine assessment and is therefore not available. Hence, we are not able to provide this additional analysis.

2. I suggest performing a subanalysis on the performance of the modified "NEWS + biomarkers" scores depending on the cause of admission (Infectious, cardiovascular, Pulmonary), as the authors did in Figure 1A for NEWS alone. Results for mortality prediction seems to heavily rely in the infectious disease group....The authors should discuss their results depending on these causes of admission.

Reply: thank you for your comment. We now performed ROC analyses in subgroups as suggested by the reviewer (stratified by main admission diagnoses, similar as shown in Figure A1). Results are shown in the following table. As you can see, NEWS + biomarkers perform best in cardiovascular disease.

Table. ROC analyses [AUC (95% CI)] for the outcome 30-day mortality in the total cohort and stratified by subg		
	total cohort	
		infectious disease
NEWS	0.73 (0.66 to 0.80)	0.76 (0.61 to 0.92)
WBC	0.64 (0.56 to 0.72)	0.57 (0.38 to 0.76)
PCT	0.71 (0.64 to 0.79)	0.69 (0.53 to 0.85)
ProADM	0.78 (0.73 to 0.84)	0.71 (0.58 to 0.84)
all combined	0.82 (0.77 to 0.88)	0.78 (0.65 to 0.91)
NEWS & WBC	0.74 (0.67 to 0.81)	0.77 (0.61 to 0.92)
NEWS & PCT	0.78 (0.72 to 0.84)	0.76 (0.60 to 0.92)
NEWS & proADM	0.82 (0.77 to 0.87)	0.78 (0.65 to 0.90)

3. Minor comment: please do not use so many "moreover" in the intro section...

Reply: thank you – we rephrased the text accordingly

Reviewer: 2

Reviewer Name: Signe Søvik, Senior consultant anaesthesiologist, Associate professor

Institution and Country: Dept. of Anaesthesia and Intensive Care
Akershus University hospital
Norway

and

Institute of Clinical Medicine
University of Oslo
Norway

General Considerations

Thank you for allowing me to review this manuscript. It reports findings from a large, unselected, adult medical ED population included at three tertiary centres in Switzerland, France, and the USA, respectively.
The study evaluates whether measurements of three inflammatory biomarkers (white blood cell count WBC, procalcitonin PCT, and mid-regional pro-Adrenomedullin MR-proADM) sampled upon patient arrival in the ED can improve the predictive power of NEWS scored upon patient arrival in the ED, outcomes being all-purpose 30-day mortality and ICU admission at any time during the hospital stay. The biomarkers had been prospectively sampled in a large study (TRIAGE), comprising 7132 patients. This study presents findings from a subsample of 1303 patients where sufficient data was available to retrospectively calculate a NEWS score.
The study is well designed, succinctly presented, and the findings seem robust. Inherent limitations are mostly well addressed. Below are some comments and suggestions that I hope may contribute to further improve the manuscript.

Specific points

Title

1. Inflammatory markers were sampled prospectively, but NEWS components were retrospectively collected, leading to a loss of almost 82% of the original patient population. I therefore strongly suggest removing the term “prospective” from the title.

Reply: thank you for your comment. As suggested, we removed the term “prospective” from the title: *“*Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multi-national, observational study”

Abstract

2. Page 3 Line 54-56:
   I suggest somehow rewriting the sentence “Combining the three inflammatory markers with NEWS improved risk stratification with regard to ICU admission compared to NEWS alone from AUC 0.70 vs. 0.65. “

   e.g.
   “Compared to using NEWS alone, combining the three inflammatory markers with NEWS improved prediction of ICU admission, AUC increasing from 0.65 to 0.70.”
   or
   “Combining the three inflammatory markers with NEWS improved prediction of ICU admission (AUC 0.70 vs. 0.65 when using NEWS alone).”

Reply: thank you for pointing this out. We changed the sentence as suggested: *"Combining the three inflammatory markers with NEWS improved prediction of ICU admission (AUC 0.70 vs. 0.65 when using NEWS alone)"*

3. Page 3 Line 50 and onwards
   See comments below regarding statistical tests assessing whether AUC values are significantly different

Reply: thank you – please see our replies below.

Introduction

4. Page 6 Line 26:
   Abbott, Ref 12, does not seem to have studied lactate. Is this an erroneous citation?

Reply: we thank you for your comment. It is indeed an erroneous citation of a different study of the same author. We corrected the citation to the following:
"Abbott TEF, Torrance HDT, Cron N, et al. A single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. European journal of internal medicine 2016;35:78-82 PubMed . doi: 10.1016/j.ejim.2016.06.014
[PubMed published Online First: 2016/06/28]"

Methods

5. Page 10 Line 13:
   The applied NEWS categories (low ≤4, moderate 5–6 and high ≥7) are the ones recommended by the Royal College of Physicians. However, in the present study data on use of supplemental oxygen (worth 2 points) were unavailable. Thus many of the patients in reality would have received a 2 point higher NEWS score. It is likely that the fraction of patients that were under-rated by 2 points due to missing information on oxygen supplements would increase from the Low via the Moderate to the High NEWS group. This could represent at systematical bias where sicker patients (probably with poorer outcomes) were assigned too low NEWS values, possibly resulting in statistical inflation of the effect of the lower NEWS values.

   I do not think this limitation distorts the study's overall findings, i.e., obvious "dose-response" effects of NEWS. Also, the authors discuss this matter towards the end of the paper. However, I recommend underlining these facts and their potential effects a little stronger in Methods.

Reply: we thank you for pointing this out. To underline this fact in the methods section we added the following statement: *"As data on supplemental oxygen was not available, results in this paper correspond to a NEWS – potentially minus 2 points."*
We further added more text to the limitation section as suggested by the reviewer: *"It is likely that the fraction of patients that were under-rated by 2 points due to missing information on oxygen supplements would increase from the low via the moderate to the high-risk NEWS group. This could represent at systematical bias where sicker patients (probably with worse outcomes) would potentially be misclassified with fewer points in the NEWS score , possibly resulting in statistical inflation of the effect of the lower NEWS values."*

   Footnotes to figures and tables should mention that NEWS in this study was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2". In its present form, the text does not distinguish between actual NEWS values reported from other studies and the present study's modified NEWS calculations.

Reply: thank you – we now mention that *"NEWS was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2""* in footnotes to tables and figures.

6. Page 11 Line 47:
   I was unable to find mean and SD being used anywhere in the text. To be deleted?

Reply: thank you – we deleted the respective term in the manuscript.

7. Page 11 Line 49:
   Replace "interquartile range (IQR)" with "quartiles" or "25th and 75th percentiles" throughout manuscript and tables. IQR denotes a single number describing the _distance between_ the 25th and the 75th percentile.

Reply: thank you for pointing this out. We replaced "interquartile range (IQR)" with "quartiles" throughout the manuscript and tables.

8. Page 11 Line 49:
   What statistical tests were used to assess group differences referred with p values in Table 1? Please add information in Methods and as footnote to table.

Reply: thank you for your comment. We added the following statement in the methods section and as footnote in table1: *"To assess group differences we used Kruskal-Wallis test for continuous, skew variables, and Pearson's chi-squared test for categorial and binary variables."*

9. Page 11 Line 56:
   Was age represented as a linear covariate or binned as categories? Describe briefly how main diagnoses and comorbidities were coded and used in the statistical models.

Reply: we thank you for pointing this out. We added a brief description in the methods section: *"Age was used as a linear covariate. According to the main admission diagnosis the following diagnostic groups were generated: Infectious disease, cardiovascular disease, metabolic disorder, malignant disease, neurological disease, gastrointestinal disease, pulmonary disease, and other disease. Comorbidities were assigned using patients' medical history and ICD-10 diagnostic codes and include chronic obstructive lung disease, heart failure, coronary heart disease, diabetes, hypertension, stroke, malignant disease and renal failure. In statistical models, comorbidities were coded as binary variables."*

10. Page 12 Line 22:
    I suggest that you add a statistical evaluation of whether AUC values from the various models did in fact differ from each other. I believe STATA has routines for such calculations.

Reply: thank you for your comment. As suggested by the reviewer we now performed statistical tests (Pearson's chi-squared test) to assess whether AUC values from the various models did statistically differ from each other. According to this we added the following statement in the methods section: *"We used Pearson's chi-squared test to compare areas under the receiver operating curve."*

Furthermore, we added p-values in the abstract, in the results section and in table 3 (changes are highlighted in bold letters).

Table 3: Discriminative performance of NEWS and biomarkers for the prediction of primary and secondary outcomes

	AUROC (95% CI)			
	30-day mortality	**p-value**	ICU admission	**p-value**
NEWS	0.73 (0.66 to 0.80)		0.65 (0.61 to 0.70)	
WBC	0.64 (0.56 to 0.72)		0.54 (0.49 to 0.59)	
PCT	0.71 (0.64 to 0.79)		0.62 (0.57 to 0.67)	
ProADM	0.78 (0.73 to 0.84)		0.67 (0.62 to 0.72)	
all combined	0.82 (0.77 to 0.88)	**0.002**	0.70 (0.65 to 0.75)	**0.006**
NEWS & WBC	0.74 (0.67 to 0.81)	**0.196**	0.65 (0.60 to 0.70)	**0.792**
NEWS & PCT	0.78 (0.72 to 0.84)	**0.004**	0.68 (0.64 to 0.73)	**0.017**
NEWS & proADM	0.82 (0.77 to 0.87)	**0.002**	0.70 (0.65 to 0.74)	**0.009**

CI, confidence interval; ICU, intensive care unit; NEWS, national early warning score; PCT, procalcitonin; MR-proADM, midregional Pro-Adrenomedullin; WBC, white blood cell count.
NEWS was calculated without oxygen supplementation data and thus represents “NEWS - potentially minus 2”
P-values correspond to the AUCs of the respective models compared to the AUC of NEWS alone and were assessed using Pearson’s chi-squared test.

Results

11. Page 13 Line 8-11:
    I suggest adding
    “.... complete information for calculation of NEWS (excluding data on supplemental oxygen)....”

Reply: thank you – we changed the respective passage exactly as suggested:
“.... complete information for calculation of NEWS (excluding data on supplemental oxygen)....”

12. What was the number of included patients from each of the three study sites, originally and in the present study?

Reply: thank you for this important comment. We added the numbers of included patients from each of the three hospitals in the results section:
“Of a total of 7,132 patients presenting to the EDs of the participating hospitals (1,000 Clearwater, 1,553 Paris, 4,579 Aarau), 1,303 (940 Clearwater, 355 Paris, 8 Aarau) patients had complete information…”

13. Page 13 Line 24:
    Does this AUC value (0.73) refer to the univariate model where NEWS alone predicts 30-day mortality, or the fully adjusted multivariate model where NEWS together with age, sex, main diagnosis, and comorbidities predict 30-day mortality (referred to in the previous sentence)?

Reply: thank you – the AUC refers to the univariate model where NEWS alone predicts 30-day mortality. Please find our detailed reply below.

It might be useful to state in Methods (e.g., Page 12 Line 17) if presented AUC calculations systematically were calculated for models fully adjusted for age, sex, main diagnosis, and comorbidities, or systematically were calculated for models only including NEWS +/- inflammatory markers. The latter is how I understand the text.

Reply: thank you for pointing this out. The AUCs presented in this manuscript were systematically calculated for models only including NEWS +/- inflammatory markers. Hence, we includedthe following statement in Methods: *"The AUCs were systematically calculated for univariate models including NEWS and/or inflammatory markers and not for models adjusted for the afore-mentioned confounders."*

14. Page 17 Line 16 and Table 3:
"... significantly improved the predictive value..." I suggest using a statistical test to compare AUC values. Many of the 95% CI's in Table 3 overlap, thus it is not obvious to the reader what models actually differed from each other.

Please add information on statistical comparison of AUC values to Methods (e.g., Page 12 Line 20 and on) if such tests were performed, or if they are added to the study.
If AUC values of the various models are not formally compared I suggest not using the term "significantly improved" since it implies a statistical evaluation.

Reply: we thank you for your comment. Please see our reply in point 10. We performed statistical tests (Pearson's chi-squared test) to assess whether AUC values from the various models did statistically differ from each other. Moreover, we did add a section in Methods. Regarding the particular citation we did found a statistically significant improvement. We therefore did not change the term "... significantly improved the predictive value..."

Tables

15. Table 1. Page 14:
Add a footnote on what statistical tests the p-values refer to.

Add a footnote informing readers that NEWS was calculated without supplemental oxygen data.

Reply: thank you. We added the following footnotes to table 1:
"NEWS was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2""
"To assess group differences we used Kruskal-Wallis test for continuous, skew variables, and Pearson's chi-squared test for categorial and binary variables."

16. Table 2. Page 16 Line 31:
"CRP" should probably be exchanged with "PCT and Pro-ADM"?

Add a footnote informing readers that NEWS was calculated without supplemental oxygen data.

Reply: thank you for pointing this out. We exchanged "CRP" with "PCT and MR-proADM" and added the following footnote to table 2:
"NEWS was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2""

17. Table A2. Page 34 Line 31:
"CRP" should probably be exchanged with "PCT and Pro-ADM"?

Add a footnote informing readers that NEWS was calculated without supplemental oxygen data.

Reply: thank you. We exchanged "CRP" with "PCT and MR-proADM" and added the following footnote to table A2:
"NEWS was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2""

18. Table A3. Page 35 Line 31:
"CRP" should probably be exchanged with "PCT and Pro-ADM"?

Add a footnote informing readers that NEWS was calculated without supplemental oxygen data.

Reply: thank you. We exchanged "CRP" with "PCT and MR-proADM" and added the following footnote to table A3:
"NEWS was calculated without oxygen supplementation data and thus represents "NEWS - potentially minus 2""

Discussion

19. Page 21 Line 47:
I suggest specifying in text, e.g., "In contrast, WBC results are available rapidly, and indeed PCT point-of-care tests that provide results within minutes are being developed".

Reply: thank you for your comment. We changed the text as suggested: "***In contrast, WBC results are available rapidly, and indeed PCT point-of-care tests that provide results within minutes are being developed"***

20. What would be the average time required to get results from an MR-proADM analysis, in a real-life ED setting?

ED patients may be in a dynamic situation, and NEWS would be expected to change for the better in patients responding favourably to treatment. Clinicians would likely evaluate results from biomarker assays in the light of the patient's condition at the time the test results become available, not at the time the blood sample was drawn. Confer the findings of Abbott et al. (reference [12]) on NEWS values pre-hospitally and upon arrival in the hospital. Please discuss briefly.

Reply: we thank you for pointing this out. Currently, results may take up to 30 minutes to become available. We are, however, aware that the company (BRAHMS, ThermoFisher Scientific) works on a fast turnaround MR-proADM POC test that would allow rapid measurement for earlier decision making. However, details are not yet available to the best of our knowledge. As suggested we now added this important point in the limitation as follows:
"However, right now there is no point of care test available that would allow rapid measurement of MR-proADM. Clinicians would likely evaluate results from biomarker assays in the light of the patient's condition at the time the test results become available, not at time the blood sample was drawn. As ED patients are in a dynamic situation, NEWS and the patient's condition might have changed already when results become available."

21. Page 22 Line 10:
Although inflammatory markers were collected prospectively in the original patient population, the present study was an analysis of NEWS calculated retrospectively, in a population where it had not been ensured that all clinical measurements necessary for the calculation of NEWS was performed systematically.

This resulted in a loss of almost 82% of the original patient population. Further, a modified NEWS still had to be used because data on supplemental oxygen was unavailable. Thus, I strongly suggest rewriting the term "secondary analysis of a prospective study" to "retrospective study" in the manuscript title and elsewhere.

Reply: thank you for your comment. We changed the term "retrospective study" as followed: *"First, since it is a study where NEWS was calculated retrospectively, associations between…"*
Moreover, we removed the term "prospective" from the title and elsewhere in the manuscript.

22. Page 22 Line 17:
Please elaborate. How could batch measurements of blood samples address selection bias? Was not the problem that patients where clinicians had decided (more or less on purpose) not to record physiological data necessary for a NEWS calculation could not be included (=selected)? My apologies if I misunderstand your text.

Reply: we thank you for this comment. The term might be misunderstanding in this context and not relevant regarding selection bias in this analysis. We removed it from the manuscript. In the original study we decided to later batch-measure the biomarkers. We aimed to improve initial triage in ED. Treating physicians were blinded to the biomarker results to also address selection bias.

23. Page 22 Line 43
I do agree that multinational inclusion of patients is a strength of this study, and therefore recommend that the number of included patients from each site is stated in Results.

Reply: thank you – we now state the number of included patients from each site in Results. Please see our reply to point 12.

VERSION 2 – REVIEW

REVIEWER	Jesus F Bermejo-Martin Group for Biomedical Research in Sepsis, BioSepsis, Hospital Clínico Universitario de Valladolid. Spain
REVIEW RETURNED	02-Nov-2018

GENERAL COMMENTS	The authors have addressed all my suggestions. This paper opens an interesting avenue to improve clinical scores for severity stratification in the ER, by adding biomarkers. Although no point of care devices are already available for leukocytes, ProADM or PCT evaluation, they are about to be released. In consequence, the results showed here are close to be applicated soon to the clinical practice..