PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The REDS score: a new scoring system to Risk-stratify Emergency Department suspected Sepsis- a derivation and validation study
AUTHORS	Sivayoham, Narani; Blake, Lesley; Tharimoopantavida, Shafi; Chughtai, Saad; Hussain, Adil; Cecconi, Maurizio; Rhodes, Andrew

VERSION 1 – REVIEW

REVIEWER	Daniel C. Richter Heidelberg University Hospital Dpt. of Anesthesiology and Critical Care Heidelberg, Germany
REVIEW RETURNED	25-Apr-2019

GENERAL COMMENTS	The authors provide a conclusive validation of a potentially new score for the ED setting to risk-stratify patients with suspected sepsis and high risk of death. After review of the manuscript I have only some minor comments: 1) Typo P8 L10: >>The difference is the percentage ...<< of instead ? 2) From a statistical point of view it should not be surprising that a combination of two moderately accurate scores (sMISSED & qSOFA) with lactate levels and sustained hypotension (both strong predictors for mortality in sepsis) would lead to a model with good ROC characteristics. However, it is legit to do so and the derivation of the new score is statistically consistent. 3) Given the components of the REDS one wonders why it would only predict mortality in suspected sepsis. I could imagine a lot of critical patients and scenarios in the ED where the REDS would predict high mortality as well (i.e. decompensated CHD --> pulmonary edema --> hypoxia --> lactate and low BP due to LV-dysfunction) 4) All included patients received 2L of fluid (or 30mL/kg) irrespective of the reason for low BP? From my perspective, this seems very pragmatic as we know that even in sepsis, there is a subset of patients that will not respond to a fluid bolus or even aggravate with fluid resuscitation. I would suggest other (simple) tests prior to liberal fluid administration (TTE-guided passive leg raise, pulse pressure analysis) in a prospective trail. 5) Inclusion criteria should be refined (one of the weak aspects of the study is the very low rate of patients admitted to the ICU due to

	DNAR orders and other circumstances) prior to starting prospective trails. 6) In terms of rational and judicious use of antimicrobials, it is important to point out that, to date, the REDS score only flags patients with high mortality (irrespective of the presence of infection). Future studies should try to gather and summarise data regarding infection and use of antimicrobials as well. 7) As mentioned in the manuscript, validation studies (preferably multi-centre setting and prospective design) are definitely needed. Overall, REDS appears promising and indeed a practical score for the ED we desperately need, as pointed out nicely by the authors.

REVIEWER	Amith Shetty Westmead Institute for medical research Published and currently working on similar combination scores
REVIEW RETURNED	28-Apr-2019

GENERAL COMMENTS	REDS score 0-3 but abstract states more than 5 or more than 7. REDS score 0-3 was that MISSED score positive 1, hypotension 1 and HL 1 but requires reading inot results section to understand it 0-12 that is true though - retrospective studies where intervention occurs are traditionally biased and thus a more robust method would be to include adverse outcomes or composite outcomes such as mortality + ICU or increased LOS REDS score out of 8? But qSOFA and RH have overlaps? Also increasingly evidence suggests lactate >2 with significant risk of adverse outcomes + revision of recent sepsis definitions support lower cut-off - this again becomes clear in latter sections of manuscript Strengths suggests missing data was included – was this imputed or considered normal? if the aim of the study is to help clinicians identify high risk patients then a single time point at beginning or end of ED journey would sound more viable - though it sounds counter-intuitive to need a score if patient has refractory hypotension or high lactate? Age and SBP were significantly different in the derivation and validation cohorts - these are two components of missed score and a key component of the qSOFA score - should be further explained in detail in limitations rather than ‘ unlikely of clinical significance? It only becomes clear towards the results section that REDS score is actually 0-12 based on all the components. This needs to occur much earlier in the methods and also in abstract – REDS score was developed with points for sMISSED, qSOFA, RH (0-3) and HL (0-3) – giving a total of 12 points? A very complicated manuscript presenting a simple concept – can we add components from existing scores and risk factors into a combination score. REDS score = sMISSED + qSOFA +

	hypotension (0-3) + lactate levels (0-3) – composite score of 12, cut-off ≥3 decent AUROC and sensitivity Manuscript could do with significant rewrite and presentation - stating REDS score is a combination of variables from sMISSED, qSOFA and also lactate (not just high lactate) and hypotension (not just refractory?) and state cut-offs for each The reviewer also provided a marked copy with additional comments. Please contact the publisher for full details.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Daniel C. Richter
Institution and Country: Heidelberg University Hospital, Dept. of Anesthesiology and Critical Care, Heidelberg, Germany Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below The authors provide a conclusive validation of a potentially new score for the ED setting to risk-stratify patients with suspected sepsis and high risk of death. After review of the manuscript I have only some minor comments:

1) Typo P8 L10: >>The difference <u>is</u> the percentage ...<< <u>of</u> instead ?
Thank you for highlighting this. This has now been corrected. The line reads as follows: The difference in the percentage of the patients admitted to the ICU is likely to reflect seasonality.

2) From a statistical point of view it should not be surprising that a combination of two moderately accurate scores (sMISSED & qSOFA) with lactate levels and sustained hypotension (both strong predictors for mortality in sepsis) would lead to a model with good ROC characteristics. However, it is legit to do so and the derivation of the new score is statistically consistent. Thank you.

3) Given the components of the REDS one wonders why it would only predict mortality in suspected sepsis. I could imagine a lot of critical patients and scenarios in the ED where the REDS would predict high mortality as well (i.e. decompensated CHD --> pulmonary edema --> hypoxia --> lactate and low BP due to LV-dysfunction). Thank you. Whilst there is no reason it would not work in other critically ill patients, at present we cannot confirm this as we have not formally studied it in other critically ill patients.

4) All included patients received 2L of fluid (or 30mL/kg) irrespective of the reason for low BP? Thank you for pointing this out and apologies for the lack of clarity. All patients did not receive 2l of fluid. We have now corrected the statement in the methods section clarifying this: 'The data collected was as follows: date of admission, method of arrival, initial vital signs [RR, HR, blood pressure (BP), oxygen saturations, temperature, blood glucose, altered mental status], final BP, initial lactate, serum albumin, INR, the use of warfarin or directly-acting oral anticoagulants (DOAC), white cell count, in those patients who received a minimum 2l of fluids the BP and lactate after the second litre of fluid, the ability to live independently,….'.
From my perspective, this seems very pragmatic as we know that even in sepsis, there is a subset of patients that will not respond to a fluid bolus or even aggravate with fluid resuscitation. I would suggest other (simple) tests prior to liberal fluid administration (TTE-guided passive leg raise, pulse

pressure analysis) in a prospective trail. Thank you for this suggestion. We will keep it in mind when designing prospective studies in the future.

5) Inclusion criteria should be refined (one of the weak aspects of the study is the very low rate of patients admitted to the ICU due to DNAR orders and other circumstances) prior to starting prospective trails. Thank you for this. We will certainly keep this in mind when designing prospective studies. However, we do feel that inclusion of the said patients may be a strength of the study as it allows us to study the natural history of the score where patients may not had advanced intervention.

6) In terms of rational and judicious use of antimicrobials, it is important to point out that, to date, the REDS score only flags patients with high mortality (irrespective of the presence of infection). Future studies should try to gather and summarise data regarding infection and use of antimicrobials as well. Again we are very grateful for your helpful suggestions which we will keep in mind when designing future studies.

7) As mentioned in the manuscript, validation studies (preferably multi-centre setting and prospective design) are definitely needed. We wholeheartedly agree.

Overall, REDS appears promising and indeed a practical score for the ED we desperately need, as pointed out nicely by the authors. Thank you!

Reviewer: 2
Reviewer Name: Amith Shetty
Institution and Country: Westmead Institute for medical research
Please state any competing interests or state 'None declared': Published and currently working on similar combination scores

Please leave your comments for the authors below:
1) REDS score 0-3 but abstract states more than 5 or more than 7. REDS score 0-3 was that MISSED score positive 1, hypotension 1 and HL 1 but requires reading in to results section to understand it 0-12
Thank you for highlighting this. We have now reworded the abstract to clarify this.

2)that is true though - retrospective studies where intervention occurs are traditionally biased and thus a more robust method would be to include adverse outcomes or composite outcomes such as mortality + ICU or increased LOS
We did not include admission to ICU as an outcome because in our institution patients may be admitted to ICU for logistical reasons. For example, a lack of sufficiently trained nurses to look after a patient with a long-term tracheostomy may be admitted to ICU rather than a ward. This is explained as one of the limitations of our study.

3)REDS score out of 8? But qSOFA and RH have overlaps? Also increasingly evidence suggests lactate >2 with significant risk of adverse outcomes + revision of recent sepsis definitions support lower cut-off - this again becomes clear in latter sections of manuscript
At first glance it may seem that the qSOFA score and RH overlap. However, through the paper we have stated that the parameters for qSOFA are taken from the initial vital signs and RH is after 2l of IV fluids. We agree that a lactate >2mmol/l is a moderate predictor of mortality. We have therefore given a lactate of 2.1-3.9 a score of 1.

4)Strengths suggests missing data was included – was this imputed or considered normal?

As stated in the methods section under the subheading of 'Sample size and missing variables' we have already stated that 'Missing variables were assumed to be normal'.

5)if the aim of the study is to help clinicians identify high risk patients then a single time point at beginning or end of ED journey would sound more viable - though it sounds counter-intuitive to need a score if patient has refractory hypotension or high lactate?
In clinical practice this score should be used when all variables have been measured. The score can be started as the patient arrives in the ED but completed when all the results are available. We have now clarified this in the penultimate paragraph of the discussion.
It is our experience that patients with refractory hypotension are not a homogenous group. Patients with RH and a REDS score 11 or 12 have a 100% mortality whereas those with a lower REDS score have a much reduced mortality rate.

6)Age and SBP were significantly different in the derivation and validation cohorts - these are two components of missed score and a key component of the qSOFA score - should be further explained in detail in limitations rather than ' unlikely of clinical significance?
Age is a component of the sMISSED score and SBP is a component of the qSOFA score. Despite these being different in the derivation and validation populations, a REDS score of ≥3 continued to identify 85% or deaths. This is a strength rather than a weakness of the score. We have now removed the above statement and added it a strength of the study in the second paragraph of the Discussion.

7)It only becomes clear towards the results section that REDS score is actually 0-12 based on all the components. This needs to occur much earlier in the methods and also in abstract – REDS score was developed with points for sMISSED, qSOFA, RH (0-3) and HL (0-3) – giving a total of 12 points?
Thank you for highlighting this. The revision of the abstract in the light of your first point should make this clearer. In addition, we have now highlighted this in the first paragraph of the discussion.

8)A very complicated manuscript presenting a simple concept – can we add components from existing scores and risk factors into a combination score. REDS score = sMISSED + qSOFA + hypotension (0-3) + lactate levels (0-3) – composite score of 12, cut-off ≥3 decent AUROC and sensitivity
Thank you. I hope we have satisfied this comment in our response to comment 1 and 7. We have changed the first paragraph of the discussion to include a statement that the REDS score is a combination of the other scores.

9)Manuscript could do with significant rewrite and presentation - stating REDS score is a combination of variables from sMISSED, qSOFA and also lactate (not just high lactate) and hypotension (not just refractory?) and state cut-offs for each
Thank you. We have re-organised and re-written some sections of the paper and hope the changes we have made are satisfactory. We have changed the term 'Hyperlactaemia' to 'Lactate' but have maintained the use of the term high lactate (HL) where we describe a lactate of ≥4mmol/l. The cut-offs for each component score is already stated in Table 3.

VERSION 2 – REVIEW

REVIEWER	Daniel C. Richter Heidelberg University Hospital, Dept. of Anesthesiology and Intensive Care Medicine, Heidelberg, Germany
REVIEW RETURNED	10-Jul-2019

GENERAL COMMENTS	Thank you for the revision of the manuscript and clarifying some points. I am looking forward to validiation studies which will hopefully support the your findings.

REVIEWER	Amith Shetty Westmead Institute for Medical Research Sydney Australia
REVIEW RETURNED	15-Jul-2019

GENERAL COMMENTS	The Authors have addressed most concerns raised except two major statistical issues: A. is REDS score a suitable screening tool - based on analysis presented , it is definitely better than qSOFA and decent performance statictics. As authors have highlighted in methods - comparing sensitivity is key and this should highlight in abstract B. The ongoing obsession of comparing AUROC for screening tests continues - please report sample size for AUROC comparison - available on MEDCALC which authors have used - https://www.medcalc.org/manual/sampling_ROC2.php https://www.ncbi.nlm.nih.gov/pubmed/6878708 the study is not powered enough for AUROC comparisons but the authors need not bother as they should be primarily comparing sensitivities for screening scores? These issues could be adequately addressed in a short paragraph in limitations Otherwise a commendable study and worthy of publication

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Daniel C. Richter
Institution and Country: Heidelberg University Hospital, Dept. of Anesthesiology and Intensive Care Medicine, Heidelberg, Germany Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below Thank you for the revision of the manuscript and clarifying some points. I am looking forward to validation studies which will hopefully support your findings.

Response: Thank you. We are confident that external validation studies will support our findings.

Reviewer: 2
Reviewer Name: Amith Shetty
Institution and Country:
Westmead Institute for Medical Research
Sydney Australia
Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below The Authors have addressed most concerns raised except two major statistical issues:

A. is REDS score a suitable screening tool - based on analysis presented , it is definitely better than qSOFA and decent performance statistics. As authors have highlighted in methods -comparing sensitivity is key and this should highlight in abstract

Response: Thank you. This is already stated in the penultimate sentence of the methods section of the abstract.

B. The ongoing obsession of comparing AUROC for screening tests continues - please report sample size for AUROC comparison - available on MEDCALC which authors have used -
https://www.medcalc.org/manual/sampling_ROC2.php
https://www.ncbi.nlm.nih.gov/pubmed/6878708
the study is not powered enough for AUROC comparisons but the authors need not bother as they should be primarily comparing sensitivities for screening scores?
These issues could be adequately addressed in a short paragraph in limitations

Response: Thank you. We have added the following sentence to the end of the paragraph that focuses on the limitations of the study, the third paragraph of the discussion section: 'Finally, we did not perform a sample size calculation to compare the AUROC curves. However, it is clear that the sensitivity for mortality of the REDS score above the cut-off point is greater than the sensitivities of the component scores above their respective cut-off points.'

Otherwise a commendable study and worthy of publication:
Response:Thank you

VERSION 3 – REVIEW

REVIEWER	Amith Shetty Westmead Institute for Medical Research
REVIEW RETURNED	01-Aug-2019

GENERAL COMMENTS	Thanks for including most if not all recommendations. I will let the readers assess the further limitations for themselves. Congratulations on your publication