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 National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: Results from a 6 year prospective cohort study
AUTHORS	Sbiti-Rohr, Diana; Kutz, Alexander; Christ-Crain, Mirjam; Thomann, Robert; Zimmerli, Werner; Hoess, Claus; Henzen, Christoph; Mueller, Beat; Schuetz, Philipp

VERSION 1 - REVIEW

REVIEWER	Mitsuhiro Sunohara University of Southern California, USA
REVIEW RETURNED	16-Jan-2016

GENERAL COMMENTS	Thank you for giving me an opportunity to review this paper. The authors' hypothesis that the NEWS score which is widely used for all the patients can be as good a predictor for patients with CAP as PSI or CURB-65, which are originally designed for CAP patients is really interesting and clinically important. I think this paper is worth accepted after several major and minor revisions shown below are performed. Major revision 1. Table 1 shows quite a lot of inconsistency with the NEWS score and PSI or CURB-65 score, i.e. low NEWS score with high PSI or CURB-65 score and vice versa. It is surprising to me that 110 patients (about a quarter) were classified as NEWS category 3 out of 459 patients with CURB-65 class 0 or 1, because CURB-65 class 0 or 1 means outpatient treatment is recommended. It is important to characterize the patients who show this inconsistency because this explains the difference of NEWS sore and PSI or CURB-65. Considering NEWS score together with PSI or CURB-65 will be beneficial for such patients, and also this could be helpful in future to design new items for scoring severity for CAP. It would be great if this paper can also describe why such patients were underestimated in CURB-65 by comparing cut off point of each item of NEWS and CURB-65 score and by mentioning the points for each item. Minor revision 1. Please mention clearly whether outpatients with CAP are included in this study. PSI and CURB-65 include outpatients, so they are used in a clinic to determine if a patient with CAP needs hospitalization. In P7L56,

	there is a sentence," For all patients, PSI, CURB-65 scores and NEWS were calculated upon admission." So I wonder if this study includes only inpatients. If this study includes only inpatients, it would be better to add another paragraph to assess the meaning. Minor revision 2. It would be nice if the percentage of the enrolled patients out of all the adult CAP patients was shown, because this is important when considering the generalization. Showing the number of patients who were excluded due to each exclusion criteria is preferable. Minor revision 3. Some patients suffer from pneumonia repeatedly especially they are terminally ill. Please mention clearly how this study treated patients who suffered from CAP more than twice. Minor revision 4. P16 LL7-12 Our analyses reveal a strong predictive value for 30-day ICU-admission and complications (empyema), even superior to the PSI and CURB-65 scores, using the NEWS. No statistical significance was observed between NEWS and PSI for ICU admission. So it would be recommended to rephrase this sentence. Minor revision 5 P17 LL44-45 NEWS has been recommended to be used not only in the initial setting but also as a trigger score for patient deterioration during hospital stay. Citation needed. Minor revision 6 Title The National Early Warning Score (NEWS) reliably improves adverse clinical outcome prediction in community-acquired pneumonia Results from a 6 year follow-up I admit that 6 year follow-up is a laborious work, but NEWS did not show the superiority in 6-year mortality rate. So in order to avoid the misunderstanding, it would be better to reconsider the title based on the conclusion.

REVIEWER	Tom Abbott Queen Mary University of London, UK
REVIEW RETURNED	01-Feb-2016

GENERAL COMMENTS	Thank you for the opportunity to review this interesting paper. This is an important area of study and provides further evaluation of NEWS in a novel clinical group. However, I think the paper would benefit from some improvement. Main concerns: 1. The paper lacks a clear hypothesis, which left me uncertain of whether the research question had been answered.

2. Consequently, the statistical analysis does not address the aims of the study very clearly.
3. There are a large amount of interesting results. However, it is not quite clear how they all come together to answer the research question.

I suspect that if the search question/hypothesis is more clearly stated and the analysis and results refined accordingly, the paper will be much easier to read.

Specific comments:

Abstract
- The authors state that this is a secondary analysis of a prospective cohort study in the design section, but later say the data are derived from a randomised controlled trial. Please confirm the study design.
- In the main outcome measure section, when describing the secondary outcome measures the authors say 'all within 30 days'. Please specify within 30 days of what?
- I am aware there is a recommended format for BMJ structured abstracts. However, it would be helpful to have a little more detail about the methods in the abstract. For example the authors use area under receiver operator characteristic curve analysis, but do not explain this method before the results and do not explain the abbreviation 'AUC' in the abstract. A little more detail would be helpful.
- Please define all abbreviations the first time the abbreviation is used
- It is not clear when was NEWS calculated? I presume on admission.

Strengths and Limitations
- Reference is made to pre-hospital patients. Did the sample include/recruit patients pre-hospital?
- 'NEWS improves prediction…' I suggest adding a comparator here
- 'impairing reproducibility to other countries' Consider re-wording, do the authors mean 'generalisability'?

Introduction
- 'We hypothesised…..' (pg 5 ln 8) – The clarity of the hypothesis could be improved. I am not sure what 'NEWS may improve these scores' means. Do the authors mean 'improve risk prediction compared to existing scores?' I suggest re-wording.

Study design
- The first line of this section describes a different design compared to the abstract. Please align with the abstract.
- Please define PCT (pg 5, ln 30).

Study design
- Is this the Center for Disease Control definition of pneumonia? If so please reference. If not, it may be helpful to include a reference for the definition the authors used.
- Pg 6 ln17 'sever' typo.
- Pg 6 ln 26. The authors mention Pro-HOSP. I assume that this is the name of the parent randomised controlled trial, but this is not clear.
- Pg 6 ln 51. I am not an expert in statistics. However, I am unsure whether classing cases with missing outcome data as survivors is appropriate. Did the authors perform a sensitivity analysis where

	these cases were considered missing and excluded from the analysis? This could introduce selection bias. - Pg 6 ln 56 onwards. The section detailing how PSI, CURB and NEWS were calculated may benefit from reference to a table or figure showing this. If there is no remaining figure/table budget then this could be included in a supplement. - Continuous data were presented as median with IQR. Were any of the data normally distributed? I might expect pack-year history of smoking, respiratory rate and body temperature to be normally distributed. - Secondary outcomes. 'CAP-associated complications (empyema)' does this only include empyema or other respiratory complications? If so, please give further detail. - Pg 8 ln8. What type of regression analyses? Given the output of analysis was odds ratios, I assume this was logistic regression – a little more detail would be helpful. - Please provide references for your choice of covariates in the multivariable regression model. - Pg 8 ln 32. Sentence starting 'Finally, we also investigated…….' I am not sure what this sentence means. I suggest re-wording. If NEWS were combined with other scores, please give further detail about how they were combined – simple addition? - I think the statistical methods could be written a little more clearly – that the authors are using these statistical tests to compare PSI, CURB and NEWS could be more prominent. Results - Pg 9 ln 13. 'There were 349….' Which NEWS categories? The categories have not been defined. This reads more like a table/figure legend. - Page 11 ln26. I disagree that AUC 0.6 show moderate predictive ability, I would consider this poor [MacDonald et al BJA 2015 vol 114(4) 598-604. Table 1 - Suggest including the definition of the NEWS categories in the legend. Discussion - pg 17 ln37 – Do the authors mean generalizability rather than reproducibility? - Pg 17 ln 47 starting 'As we disposed…' I'm not sure what this sentence means, suggest re-wording.

REVIEWER	Daniel Silcock NHS Glasgow and Clyde, UK
REVIEW RETURNED	17-Feb-2016

GENERAL COMMENTS	The current study aims to identify whether NEWS is of prognostic significance in patients diagnosed with community acquired pneumonia However, I have a few issues with the study as submitted. Mainly, it is not clear to me whether this is a new study using the historical data or a continuation of the original ProHosp study - the trial registration quoted suggests the latter. The methods section does not make clear what was done in the initial study and what has been done subsequently. Presumably the 6 year follow up is new data. It

	is not clear whether ethical approval was sought or required for this additional data collection as it is not part of the original study design as registered on controlled-trials.com The strengths and limitations section seems to mostly list strengths of the NEWS rather than the study itself, which is what I would expect to find in this section The introduction section opens with 'Today, it is recommended...' but then cites a paper from 1985 - a more up-to-date reference would I think be preferable The methods section does not clearly differentiate between the original study and the current work. In particular, I presume that NEWS scores have been calculated retrospectively on admission data, and so the line 'For all patients.....calculated upon admission' is misleading. Regarding secondary outcomes, are you defining CAP-associated complications as solely empyema? If not then a description of what are considered complications would be helpful. Tables 2/3 and related text quote odds ratios but do not specify what NEWS cut-off is being investigated (low vs med/high, low/med vs high, or low vs high?). Clarification on how the NEWS and CURB65/PSI scores are combined would be helpful - is this simple addition of the base scores or is allowance made for the differing denominators? Conclusions seem reasonable based on the supplied results There seem to be two competing interest statements, again appearing to separate the old from the new studies

REVIEWER	Richard Costello MD Associate Professor of Medicine Royal College of Surgeons in Ireland Dublin Ireland
REVIEW RETURNED	26-Apr-2016

GENERAL COMMENTS	The National Early Warning Score (NEWS) reliably improves adverse clinical outcome prediction in community-acquired pneumonia: Results from a 6 year follow-up In this paper Sbiti-Rohr and colleagues compare the NEWS categories of scores with outcomes in patients admitted to Hospitals in Switzerland who were enrolled in a clinical trial. The results of the study indicate that the score on arrival to Hospital has good predictive validity for subsequent clinically meaningful outcomes including admission to ICU, and the development of empyema and a little less intuitively to long term survival. Importantly, both nursing home and community patients were included. Some comments and further analysis: 1. Why were hospital patients who developed a healthcare setting pneumonia excluded and similarly the other exclusion criteria? 2. Who made the decision to transfer the patient to ICU and was the NEWS score used as part of the decision making process? 3. What domains within NEWS were most valid in predicting the admission to ICU, and mortality? 4. What was the time difference in time between admission to

	Hospital and the transfer to ICU? Was information on the subsequent days NEWS available, or change in NEWS and does this help in further refining the validity of the score? 5. Were other clinically meaningful measures of complications of pneumonia such as septicemia, ward based non invasive ventilation, high flow oxygen for respiratory shunt or the development of SIADH (syndrome of inappropriate ADH release with hyponatremia), acute renal failure or delirium recorded and are they also distributed across the severity of NEWS? 6. What is the mechanism of acute physiological changes impacting on long term survival? I am surprised by the long term association of higher scores and long term mortality. Is the higher NEWS not simply a reflection that some patients are susceptibile to acquire pneumonia and that these weaker patients have an increased mortality on follow up or is there an alternative i.e. that the increased long term mortality is a consequence of such a dramatic host response and/or ICU or empyema? 7. Can you comment on the re-admission rates and what you know about the cause of re-admission in the study cohort. It is interesting that the sickest are not necessarily the ones who are re-admitted, so is this a reflection of some variation in quality of healthcare? Was quality of care measured in the study? 8. Were patients who had a pneumonia but a score of 3 or less always admitted to the Hospital, in which case this may have some bias. 9. Can you explain how you developed the model of combined NEWS and PSI and CURB-65. 10. Was the NEWS calculated before oxygen was started or while on O2? Minor Comments Spelling and grammar in particular in the discussion need to be corrected, ive up loaded the paper

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name
Mitsuhiro Sunohara

The authors' hypothesis that the NEWS score which is widely used for all the patients can be as good a predictor for patients with CAP as PSI or CURB-65, which are originally designed for CAP patients is really interesting and clinically important.
I think this paper is worth accepted after several major and minor revisions shown below are performed.

Major revision 1.
Table 1 shows quite a lot of inconsistency with the NEWS score and PSI or CURB-65 score, i.e. low NEWS score with high PSI or CURB-65 score and vice versa. It is surprising to me that 110 patients (about a quarter) were classified as NEWS category 3 out of 459 patients with CURB-65 class 0 or 1, because CURB-65 class 0 or 1 means outpatient treatment is recommended.
It is important to characterize the patients who show this inconsistency because this explains the difference of NEWS sore and PSI or CURB-65. Considering NEWS score together with PSI or CURB-65 will be beneficial for such patients, and also this could be helpful in future to design new items for scoring severity for CAP.
It would be great if this paper can also describe why such patients were underestimated in CURB-65 by comparing cut off point of each item of NEWS and CURB-65 score and by mentioning the points

for each item.
REPLY: We agree that this is quite surprising, although PSI and CURB are mortality scores, while NEWS is rather a severity score and thus based on different parameters. While PSI and CURB65 are more age dominated – NEWS is more vital sign dominated. Not surprisingly, the patients misclassified with PSI / CURB65 (i.e. patients needing ICU admission despite low scores) were younger patients with low burden of comorbidities, but abnormal vital signs. We discuss this point as follows in the result and discussion sections
• "Patients that were misclassified by the PSI score as low risk (PSI class 1 or 2) but correctly identified by NEWS had a younger age (median age 49 years vs 74 years), less comorbidities (heart and renal failure, coronary heart disease) and more frequent deterioration (chills, oxygenation) of vital signs compared to patients that were correctly identified by both scores."
• "Interestingly, the PSI contains very similar physiological parameters as used for NEWS calculation. Still, NEWS was superior for adverse outcome prediction but inferior in regard to mortality prediction. This may be explained by the fact that PSI is age-dominated and while age is a good predictor for mortality, aged people at the end of life may be less often admitted to the ICU. NEWS sets the main focus on the acute condition (e.g. need for supplemental oxygen or altered level of consciousness) allowing better evaluation for the eventual of need for ICU-admission. Interestingly, in line with this, we found that younger patients with lower burden of comorbidities and more severe deterioration of vital signs were at higher risk for being misclassified as "low risk" with PSI but correctly identified with NEWS. This patient population may thus show the most benefit of combination of both scores..".

Minor revision 1.
Please mention clearly whether outpatients with CAP are included in this study.
REPLY: We have now changed this accordingly and now state that: "Inpatients and outpatients were eligible for the study." (methods) and "Most patients were treated as inpatients with 8.8% patients being treated on an outpatient basis." (results).

PSI and CURB-65 include outpatients, so they are used in a clinic to determine if a patient with CAP needs hospitalization. In P7L56, there is a sentence," For all patients, PSI, CURB-65 scores and NEWS were calculated upon admission." So I wonder if this study includes only inpatients.
REPLY: We have now changed this to: "For all patients, PSI, CURB-65 scores and NEWS were calculated upon admission to the emergency department".

If this study includes only inpatients, it would be better to add another paragraph to assess the meaning.
REPLY: please see above

Minor revision 2.
It would be nice if the percentage of the enrolled patients out of all the adult CAP patients was shown, because this is important when considering the generalization. Showing the number of patients who were excluded due to each exclusion criteria is preferable.
REPLY: Because this is a secondary analysis of a previous trial, these numbers have been shown previously in all detail in our reference 28 (Schuetz P, Christ-Crain M, Thomann R, et al. Effect of procalcitonin-based guidelines vs standard guidelines on antibiotic use in lower respiratory tract infections: the ProHOSP randomized controlled trial. JAMA 2009;302(10):1059-66.).
We now added a short statement in this regard as follows "As previously reported, we included 1381 out of from 1825 screened patients in the study of which 925 had CAP and were used for the current analysis 28.".

Minor revision 3.
Some patients suffer from pneumonia repeatedly especially they are terminally ill.
Please mention clearly how this study treated patients who suffered from CAP more than twice.

REPLY: We did not have an exclusion criterion based on the number of CAP episodes – all patients, however, were only included once in the study.

Minor revision 4.
P16 LL7-12
Our analyses reveal a strong predictive value for 30-day ICU-admission and complications (empyema), even superior to the PSI and CURB-65 scores, using the NEWS.
No statistical significance was observed between NEWS and PSI for ICU admission. So it would be recommended to rephrase this sentence.
REPLY: We have now changed this accordingly to: “Our analyses reveal a strong predictive value for 30-day ICU-admission and complications (empyema) using the NEWS.”.

Minor revision 5
P17 LL44-45
NEWS has been recommended to be used not only in the initial setting but also as a trigger score for patient deterioration during hospital stay.
Citation needed.
REPLY: We have now added a reference as suggested

Minor revision 6
Title
The National Early Warning Score (NEWS) reliably improves adverse clinical outcome prediction in community-acquired pneumonia Results from a 6 year follow-up
I admit that 6 year follow-up is a laborious work, but NEWS did not show the superiority in 6-year mortality rate. So in order to avoid the misunderstanding, it would be better to reconsider the title based on the conclusion.
REPLY: We have now changed this according to your and the editors comment to: “The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: Results from a 6 year prospective cohort study”.

Reviewer: 2

Reviewer Name
Tom Abbott

Main concerns:

1. The paper lacks a clear hypothesis, which left me uncertain of whether the research question had been answered.
REPLY: We have now added the following statement: “Our hypothesis was that NEWS would show an association with short and long-term adverse outcome in patients with CAP and possibly improve risk prediction compared to established risk assessment tools such as PSI and CURB-65. The aim of our study was thus to compare the accuracy of NEWS to predict mortality and adverse clinical outcomes with PSI and CURB-65 in a well characterised cohort of CAP patients. ”.

2. Consequently, the statistical analysis does not address the aims of the study very clearly.
REPLY: Please see above

3. There are a large amount of interesting results. However, it is not quite clear how they all come together to answer the research question.
I suspect that if the search question/hypothesis is more clearly stated and the analysis and results refined accordingly, the paper will be much easier to read.

REPLY: We have now reworked the manuscript to added to clarity to the analyses and results.

Specific comments:

Abstract
- The authors state that this is a secondary analysis of a prospective cohort study in the design section, but later say the data are derived from a randomised controlled trial. Please confirm the study design.
REPLY: The initial trial was a randomized trial investigating antibiotic stewardship. We also followed the patients for over 6 year of follow up to investigate prognostic parameters as done in this analysis. Because we found that the initial randomization did not have any association with outcome (it was a non-inferiority study), it was no longer considered in the analysis. We therefore describe the study as a prospective cohort study with patients included in a previous randomized trial. This has now been clarified in the manuscript.
- In the main outcome measure section, when describing the secondary outcome measures the authors say 'all within 30 days'. Please specify within 30 days of what?
REPLY: We have now changed this accordingly to: "all within 30 days after admission".
- I am aware there is a recommended format for BMJ structured abstracts. However, it would be helpful to have a little more detail about the methods in the abstract. For example the authors use area under receiver operator characteristic curve analysis, but do not explain this method before the results and do not explain the abbreviation 'AUC' in the abstract. A little more detail would be helpful.
REPLY: We have now added the following statement "We used regression models to study associations of baseline risk scores and outcomes with the area under the receiver operating curve (AUC) as a measure of discrimination.".
- Please define all abbreviations the first time the abbreviation is used
REPLY: We have now changed this throughout the manuscript
- It is not clear when was NEWS calculated? I presume on admission.
REPLY: We have now changed this accordingly to: "For all of them the NEWS, PSI and CURB-65 scores were calculated upon admission to the emergency department.".

Strengths and Limitations
- Reference is made to pre-hospital patients. Did the sample include/recruit patients pre-hospital?
REPLY: We have now changed this to: "In the emergency department setting, NEWS is an adequate tool for risk stratification.".
- 'NEWS improves prediction…' I suggest adding a comparator here
REPLY: We have now changed this to: "NEWS is a predictor for risk of ICU admission and clinical complications.".
- 'impairing reproducibility to other countries' Consider re-wording, do the authors mean 'generalisability'?
REPLY: thank you – this is now rephrased to "This study was limited to Swiss, predominantly Caucasian patients, limiting the generalizability of results"

Introduction
- 'We hypothesised…..' (pg 5 ln 8) – The clarity of the hypothesis could be improved. I am not sure what 'NEWS may improve these scores' means. Do the authors mean 'improve risk prediction compared to existing scores?' I suggest re-wording.
REPLY: We have now changed this to: "Our hypothesis was that NEWS would show an association with short and long-term adverse outcome in patients with CAP and possibly improve risk prediction compared to established risk assessment tools such as PSI and CURB-65.".

Study design
- The first line of this section describes a different design compared to the abstract. Please align with

the abstract.
REPLY: Changed as requested
- Please define PCT (pg 5, ln 30).
REPLY: Changed as requested to: “The primary aim of the study was to examine whether a procalcitonin (PCT)-guided algorithm”.

Study design
- Is this the Center for Disease Control definition of pneumonia? If so please reference. If not, it may be helpful to include a reference for the definition the authors used.
REPLY: We have now added that this definition is based on the American Thoracic Society guidelines.
- Pg 6 ln17 ‘sever’ typo.
REPLY: Changed as requested
- Pg 6 ln 26. The authors mention Pro-HOSP. I assume that this is the name of the parent randomised controlled trial, but this is not clear.
REPLY: Changed as requested
- Pg 6 ln 51. I am not an expert in statistics. However, I am unsure whether classing cases with missing outcome data as survivors is appropriate. Did the authors perform a sensitivity analysis where these cases were considered missing and excluded from the analysis? This could introduce selection bias.
REPLY: Censoring is a standard procedure in Cox regression where time to event is only counted as long as patients are in the study. Still, we agree that high lost to follow up could bias the analysis. In fact, we did have an excellent follow-up in this cohort with >95% of patients having up to date information about survival status.
- Continuous data were presented as median with IQR. Were any of the data normally distributed? I might expect pack-year history of smoking, respiratory rate and body temperature to be normally distributed.
REPLY: For consistency, we decided to show all data as median and IQR although some data may have shown a near normal distribution.
- Secondary outcomes. ‘CAP-associated complications (empyema)’ does this only include empyema or other respiratory complications? If so, please give further detail.
REPLY: it only includes empyema.
- Pg 8 ln8. What type of regression analyses? Given the output of analysis was odds ratios, I assume this was logistic regression – a little more detail would be helpful.
REPLY: We have now changed this accordingly to: “We used univariate and multivariate regression analyses to assess the association between the prognostic scores and the different outcomes. We report hazard ratio (HR) for all time to event analyses, and odds ratios (ORs) for all logistic regression analyses..”.
- Please provide references for your choice of covariates in the multivariable regression model.
REPLY: We pre-defined predictors based on pathophysiological considerations but cannot reference this to a previous paper. Particularly, we used age and gender as “basic covariates for our simple model (1) and age, gender as well as main comorbidities (chronic obstructive pulmonary disease [COPD], congestive heart failure, neoplastic disease, diabetes mellitus, coronary artery disease, cerebrovascular disease, peripheral artery occlusive disease [PAOD], chronic renal failure) for our more advanced model (2).
- Pg 8 ln 32. Sentence starting ‘Finally, we also investigated…….’ I am not sure what this sentence means. I suggest re-wording. If NEWS were combined with other scores, please give further detail about how they were combined – simple addition?
REPLY: We have now reworded this to: “Finally, we also investigated whether NEWS improves PSI and CURB-65 by comparing the AUC of a statistical model limited to the single CAP scores alone with a joint statistical regression model combining the CAP score and NEWS each.”.
- I think the statistical methods could be written a little more clearly – that the authors are using these

statistical tests to compare PSI, CURB and NEWS could be more prominent.
REPLY: We have now adapted the text accordingly

Results
- Pg 9 ln 13. 'There were 349….' Which NEWS categories? The categories have not been defined. This reads more like a table/figure legend.
REPLY: We have now changed this to: "Overall, we included 925 CAP patients and the median follow- up was 6.1 year. Baseline characteristics overall and according to NEWS categories low (0-4 points), medium (5-6 points) or high (≥ 7 points) risk are presented in Table 1.".
- Page 11 ln26. I disagree that AUC 0.6 show moderate predictive ability, I would consider this poor [MacDonald et al BJA 2015 vol 114(4) 598-604.
REPLY: We agree and changed this to "low".

Table 1
- Suggest including the definition of the NEWS categories in the legend.
REPLY: We have now added this as suggested

Discussion
- pg 17 ln37 – Do the authors mean generalizability rather than reproducibility?
REPLY: We have now added this as suggested
- Pg 17 ln 47 starting 'As we disposed…' I'm not sure what this sentence means, suggest re-wording.
REPLY: We have now changed this sentence to: "Because parameters for calculation of NEWS were only collected upon admission to the emergency department, no follow-up analyses were done.".

Reviewer: 3

Reviewer Name
Daniel Silcock

The current study aims to identify whether NEWS is of prognostic significance in patients diagnosed with community acquired pneumonia However, I have a few issues with the study as submitted. Mainly, it is not clear to me whether this is a new study using the historical data or a continuation of the original ProHosp study - the trial registration quoted suggests the latter. The methods section does not make clear what was done in the initial study and what has been done subsequently. Presumably the 6 year follow up is new data. It is not clear whether ethical approval was sought or required for this additional data collection as it is not part of the original study design as registered on controlled-trials.com
REPLY: We have now changed the methods to make this more clear ("This is a prospective cohort study with patients included in a prospective randomized non-inferiority trial included 925 CAP patients with a 6 years follow-up.") and also expanded the statement about ethical approval to: "All local ethical committees approved the initial study protocol, and also gave permission to do a 6-year follow-up study. All patients gave written informed consent to the initial study and the follow-up analysis including the current analysis.".

The strengths and limitations section seems to mostly list strengths of the NEWS rather than the study itself, which is what I would expect to find in this section
REPLY: We have now changed this as suggested

The introduction section opens with 'Today, it is recommended...' but then cites a paper from 1985 - a

more up-to-date reference would I think be preferable
REPLY: We have now updated the references

The methods section does not clearly differentiate between the original study and the current work. In particular, I presume that NEWS scores have been calculated retrospectively on admission data, and so the line 'For all patients.....calculated upon admission' is misleading.
REPLY: We agree and changed this to “NEWS was calculated retrospectively on admission data based on the following six physiological parameters: respiratory rate, oxygen saturation, temperature, systolic blood pressure, pulse rate and level of consciousness.”.
Regarding secondary outcomes, are you defining CAP-associated complications as solely empyema? If not then a description of what are considered complications would be helpful.
REPLY: Yes – its only empyema
Tables 2/3 and related text quote odds ratios but do not specify what NEWS cut-off is being investigated (low vs med/high, low/med vs high, or low vs high?). Clarification on how the NEWS and CURB65/PSI scores are combined would be helpful - is this simple addition of the base scores or is allowance made for the differing denominators?
REPLY: We used the score as a continuous predictor in the model and now state that odds ratios correspond to a point increase in the score (Table 2).

Conclusions seem reasonable based on the supplied results

There seem to be two competing interest statements, again appearing to separate the old from the new studies
REPLY: We have now changed this and dropped in the statement regarding the initial study

Reviewer: 4

Reviewer Name
Richard Costello MD Associate Professor of Medicine

In this paper Sbiti-Rohr and colleagues compare the NEWS categories of scores with outcomes in patients admitted to Hospitals in Switzerland who were enrolled in a clinical trial. The results of the study indicate that the score on arrival to Hospital has good predictive validity for subsequent clinically meaningful outcomes including admission to ICU, and the development of empyema and a little less intuitively to long term survival. Importantly, both nursing home and community patients were included.

Some comments and further analysis:
1. Why were hospital patients who developed a healthcare setting pneumonia excluded and similarly the other exclusion criteria?
REPLY: We wanted to focus on a more homogenous CAP patient population and thus excluded these patients

2. Who made the decision to transfer the patient to ICU and was the NEWS score used as part of the decision making process?
REPLY: The decision for ICU transfer was up to the discretion of the treating physician – who were not aware of the NEWS score. This is now stated in the revised manuscript.

3. What domains within NEWS were most valid in predicting the admission to ICU, and mortality?
REPLY: We only looked at the NEWS score as a continuous predictor – but not within all subdomains

of NEWS. Because this score is already used in many facilities, our hypothesis was more pragmatic, i.e. how well does the score – as recommended – perform in CAP patients.
4. What was the time difference in time between admission to Hospital and the transfer to ICU? Was information on the subsequent days NEWS available, or change in NEWS and does this help in further refining the validity of the score?
REPLY: Unfortunately, we only had baseline variables for calculation of NEWS, but no follow up data. This is now stated in the revised manuscript as follows: "Because parameters for calculation of NEWS were only collected upon admission to the emergency department, no follow-up analyses were done.".

5. Were other clinically meaningful measures of complications of pneumonia such as septicemia, ward based non invasive ventilation, high flow oxygen for respiratory shunt or the development of SIADH (syndrome of inappropriate ADH release with hyponatremia), acute renal failure or delirium recorded and are they also distributed across the severity of NEWS?
REPLY: Unfortunately, we did not record these types of outcome measures
6. What is the mechanism of acute physiological changes impacting on long term survival? I am surprised by the long term association of higher scores and long term mortality. Is the higher NEWS not simply a reflection that some patients are susceptibile to acquire pneumonia and that these weaker patients have an increased mortality on follow up or is there an alternative i.e. that the increased long term mortality is a consequence of such a dramatic host response and/or ICU or empyema?
REPLY: We agree that this is a possibility. The main result regarding mortality is that NEWS is associated with this outcome, but that other scores (which were derived for mortality prediction) shown better accuracy (i.e. PSI and CURB65).
7. Can you comment on the re-admission rates and what you know about the cause of re-admission in the study cohort. It is interesting that the sickest are not necessarily the ones who are re-admitted, so is this a reflection of some variation in quality of healthcare? Was quality of care measured in the study?
REPLY: We did not collect data on reasons for rehosp in this population because it was just a secondary outcome and not the main outcome of interest. Thus, unfortunately, we can not commend on this interesting issue.
8. Were patients who had a pneumonia but a score of 3 or less always admitted to the Hospital, in which case this may have some bias.
REPLY: NEWS was not available upon admission and thus did not influence admission decsions.
9. Can you explain how you developed the model of combined NEWS and PSI and CURB-65.
REPLY: We have now expanded on this and state in the statistical section that "Finally, we also investigated whether NEWS improves PSI and CURB-65 by comparing the AUC of a statistical model limited to the single CAP scores alone with a joint statistical regression model combining the CAP score and NEWS each.".
10. Was the NEWS calculated before oxygen was started or while on O2?
REPLY: oxygenation was used before start of O2 therapy

Minor Comments
Spelling and grammar in particular in the discussion need to be corrected, ive up loaded the paper
REPLY: We have now reworked the manuscript checking the spelling throughout.

VERSION 2 – REVIEW

REVIEWER	Mitsuhiro Sunohara University of Southern California, U.S.A.
REVIEW RETURNED	08-Jun-2016

GENERAL COMMENTS	Thank you for giving me an honor to review your great manuscript. I believe this paper includes important data and messages for those who are struggling to treat pneumonia.

REVIEWER	Tom Abbott Queen Mary University of London
REVIEW RETURNED	24-May-2016

GENERAL COMMENTS	This is an interesting paper and much improved from the version perviously submitted. I suggest accepting this paper for publication with further revisions. 1. The explanation of the study design is still a little confusing. I would suggest wording this either as a secondary analysis of data from a previously published randomised randomised controlled trial, or a cohort study using data from patients enrolled in a previous RCT - the current use of the word prospective implies it was a cohort study from beginning, which is inaccurate. Please accept my apologies if my previous comment was not clear. Abstract and Methods (page 6). 2. Some of the language used throughout the paper is imprecise. If the authors do not speak English as their first language, the paper may benefit from editorial input from a native English speaker, to improve the readability of the paper. 3. I disagree that AUROC of 0.73 for ICU admission shows high discriminative ability - this is moderate at best. See McDonald et al BJA 2015. I suggest changing this - Abstract and discussion (page 17) 4. Page 14. The authors mention a statistically significant relationship in the second paragraph. I suggest including either the p-value or OR/HR with confidence interval in the main text.

REVIEWER	Daniel Silcock NHS Glasgow+Clyde
REVIEW RETURNED	07-Jun-2016

GENERAL COMMENTS	Minor revision (throughout) - The paper throughout uses 'complications' when really only empyema is meant. I would suggest that other complications of pneumonia (sepsis, SIADH etc) are at least equally important and should be included, or if data is unavailable, the description should be changed to reflect that only empyema is meant. It should also be clarified whether empyema means clinically suggested empyema (eg radiographic evidence of effusion) or laboratory confirmed empyema Major revision: The method used to combine the NEWS and PSI scores is possibly now more unclear than previously, and it would not be possible to recreate the statistical model used to analyse the result based on the information provided. Additionally, frequent mentions are made of adding NEWS to PSI, suggesting arithmetical addition rather than more complicated combination. The resulting score would appear to be too complicated to be useful clinically.

	Minor revision: It is not mentioned in the text whether supplemental oxygen was included in the analysis for NEWS - this makes a difference as 2 points are allotted for supplemental oxygen use Minor revision: Spelling mistake - Strengths and weaknesses Minor revision: Better clarification needs to be made between NEWS/PSI points and NEWS/PSI categories Minor revision: It should be noted that NEWS was developed as a track and trigger system and only later validated (with lower numbers) for use as an initial assessment tool. At several points in the paper it is suggested the reverse is true (particularly "NEWS has been recommended to be used not only in the initial setting but also as a trigger score for patient deterioration during hospital stay" cited reference is the initial publication of the NEWS score) Presumably NEWS was not available at time of admission as the score was yet to be published but clarification should be made of whether treatment decisions were likely to be made on the basis of CURB/PSI (for example as part of an outpatient treatment protocol.)

REVIEWER	Richard Costello RCSI Dublin
REVIEW RETURNED	02-Jun-2016

GENERAL COMMENTS	Im happy with the responses provided.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Reviewer Name: Tom Abbott

1. The explanation of the study design is still a little confusing. I would suggest wording this either as a secondary analysis of data from a previously published randomised controlled trial, or a cohort study using data from patients enrolled in a previous RCT - the current use of the word prospective implies it was a cohort study from beginning, which is inaccurate. Please accept my apologies if my previous comment was not clear. Abstract and Methods (page 6).

REPLY: As suggested, we now reworded the text stating the study design in the Abstract and Methods as «Secondary analysis of patients included in a previous randomized controlled trial» (Abstract and Methods)

2. Some of the language used throughout the paper is imprecise. If the authors do not speak English as their first language, the paper may benefit from editorial input from a native English speaker, to improve the readability of the paper.

REPLY : We have now edited the text for better language and hope this meets the reviewer`s expectations.

3. I disagree that AUROC of 0.73 for ICU admission shows high discriminative ability - this is moderate at best. See McDonald et al BJA 2015. I suggest changing this - Abstract and discussion (page 17)

REPLY : We agree and changed this in the Abstract and Discussion to : « moderate »

4. Page 14. The authors mention a statistically significant relationship in the second paragraph. I suggest including either the p-value or OR/HR with confidence interval in the main text.

REPLY: We now added the OR with confidence interval for the unadjusted model and the p-value for the adjusted model accordingly to :
« Table 3 shows the unadjusted and adjusted regression analysis investigating the association of NEWS with adverse clinical outcomes, namely ICU-admission, complications and re-hospitalisation. The results were statistically significant for NEWS as a predictor for ICU-admission (OR 1.29 [1.2, 1.39] and complications (OR 1.16 [1.04, 1.29] within 30 days after admission. This was also true after adjustment for age, gender and comorbidities ($p<0.01$, each). Concerning re-hospitalization, no significant association was found. »

Reviewer: 3

Reviewer Name: Daniel Silcock

1. Minor revision (throughout)
The paper throughout uses 'complications' when really only empyema is meant. I would suggest that other complications of pneumonia (sepsis, SIADH etc) are at least equally important and should be included, or if data is unavailable, the description should be changed to reflect that only empyema is meant. It should also be clarified whether empyema means clinically suggested empyema (eg radiographic evidence of effusion) or laboratory confirmed empyema

REPLY : In this cohort of patients, only empyema was found as a complications while no case of ARDS was found. We now use the term empyema instead of complications as suggested. We also no in more detail explain how empyema was diagnosed: “We recorded all patients with empyema diagnosed by their treating physicians by ultrasound and laboratory examination.”

2. Major revision
The method used to combine the NEWS and PSI scores is possibly now more unclear than previously, and it would not be possible to recreate the statistical model used to analyse the result based on the information provided. Additionally, frequent mentions are made of adding NEWS to PSI, suggesting arithmetical addition rather than more complicated combination. The resulting score would appear to be too complicated to be useful clinically.

REPLY : we have now rephrased the statistical section as follows: “Finally, we also investigated whether NEWS adds prognostic information to PSI and CURB-65 in regard to discrimination. For this purpose, we compared the AUC of a regression model limited to the PSI score with a binary regression model including PSI and NEWS. The same was done for CURB65.”
We also changed the text in regard to “adding” as suggested and now use the term “combine” instead.

3. Minor revision
It is not mentioned in the text whether supplemental oxygen was included in the analysis for NEWS - this makes a difference as 2 points are allotted for supplemental oxygen use

REPLY : In the Methods, section « Assessment of vital status and score assignment », third paragraph we describe this as follows :
“NEWS was calculated retrospectively on admission data based on the following six physiological parameters: respiratory rate, oxygen saturation, temperature, systolic blood pressure, pulse rate and level of consciousness. Every continuous variable scores a maximum of 3 points, whereas the need for supplemental oxygen and the level of consciousness are binary coded with zero points if absent/normal and 2 or 3 points if present/altered respectively.”

4. Minor revision
Spelling mistake - Strengths and weaknesses

REPLY: Thank you - changed as requested

5. Minor revision
Better clarification needs to be made between NEWS/PSI points and NEWS/PSI categories

REPLY : We have now adjusted this throughout the manuscript. NEWS is now throughout the paper described in risk categories whereas the PSI is described in risk classes. We rephrased the PSI description as follows: “The PSI includes 20 variables resulting in a point score and classifies the patients with CAP into five risk…”

6. Minor revision
It should be noted that NEWS was developed as a track and trigger system and only later validated (with lower numbers) for use as an initial assessment tool. At several points in the paper it is suggested the reverse is true (particularly "NEWS has been recommended to be used not only in the initial setting but also as a trigger score for patient deterioration during hospital stay" cited reference is the initial publication of the NEWS score)

REPLY: Thank you, we have now rephrased this in the Introduction as follows: “Its purpose was to introduce a standardised trigger-system to identify acutely ill patients throughout hospitalisation”; and in the Discussion as follows: “NEWS has been primarily recommended to be used as a trigger score for patient deterioration during hospital stay and not in the initial setting”

7. Presumably NEWS was not available at time of admission as the score was yet to be published but clarification should be made of whether treatment decisions were likely to be made on the basis of CURB/PSI (for example as part of an outpatient treatment protocol.)

REPLY: The treatment decisions were not made on the basis of these scores (CURB-65, PSI)

VERSION 3 – REVIEW

REVIEWER	Tom Abbott Clinical Research Training Fellow, Queen Mary University of London, UK
REVIEW RETURNED	28-Jul-2016

GENERAL COMMENTS	This is an important topic and an interesting paper. This version of the paper is much improved compared to previous versions. The authors/editors might consider the following two (minor) points:

	1. Table 1. Add an explanation of the statistical test the p-vlaue represents to the table legend. 2. Discussion section. The authors mention potential confounding as a result of the exclusion criteria for the parent trial. Small additional detail could be added to explain how this might affect the results.

REVIEWER	Daniel Silcock NHS Glasgow and Clyde
REVIEW RETURNED	03-Aug-2016

GENERAL COMMENTS	Apologies for the delay The latest changes satisfy my concerns