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)	Comparison of SIRS criteria and qSOFA score for identifying culture-positive sepsis in the Emergency Department A prospective cross-sectional multicenter study
AUTHORS	Mignot-Evers, Lisette; Raaijmakers, Vivian; Buunk, Gerba; Brouns, Steffie; Romano, Lorenzo; van Herpt, Thijs; Gharbharan, Arvid; Dieleman, Jeanne; Haak, Harm

VERSION 1 – REVIEW

REVIEWER	Shetty, Amith Westmead Millennium Institute for Medical Research, NHMRC Centre for research excellence in Critical excellence
REVIEW RETURNED	21-Jun-2020

GENERAL COMMENTS	The authors have conducted a prospective observational study between two hospital ED which were slightly different in staffing and resourcing (not major difference) - one which uses SIRS to identify septic patients and other that uses qSOFA patients. The aim of the study is reported as need to check SIRS versus qSOFA in identifying sepsis. But in serious contradiction to recent sepsis definitions (from which qSOFA was derived) the authors go on to redefine sepsis as culture positive patients! the study becomes very convoluted to then find in figure two whether a group of patients were identified in one hospital or the other or the mixed cohort was from both hospitals ED (most likely). The study then goes on to compare 5-cohorts based on treatment received. Whilst this is very interesting analysis and suggest some obvious signals, (increased mortality in qSOFA + cohorts and higher in SIRS negative qSOFA+ patients) the study is not powered enough to suggest HR purely based on a singular criteria model is suggestive - this section requires statistical review - if the manuscript is resubmitted. Cohorts 4 and 5 have 2 and 6 patients respectively. I have not extensively reviewed the document as the above seriously jeopardizes reviewing the article further. I believe there are some serious design flaws in the study: 1. redefine sepsis as suspected infection + SOFA score ≥2 or redefine outcome as identifying culture positive patients 2. During the study period identify the number of patients who were diagnosed with sepsis or culture positive (based on changes made from point 1) and not just inclusions. Screening test performance of SIRS and qSOFA tools should occur on this final cohort of patients and sensitivity specificity and PPV/NPV needs to

	be calculated. A screening test would need very high sensitivity to be considered superior not accuracy especially in the ED. the study adds information about the significance of qSOFA in flagging patients who are at risk of adverse outcomes - but all other claims are discounted due to the definition bias introduced with defining sepsis as culture positive patients.

REVIEWER	Keijzers, Gerben Griffith Univ, School of Medicine
REVIEW RETURNED	31-Oct-2020

GENERAL COMMENTS	Thank you for asking me to review this dual site study from The Netherlands aiming to assess diagnostic accuracy of SIRS and Sepsis in the ED. Sepsis is a clinical syndrome with no tissue diagnostic or reliable serological test. Clinical features can be non-specific. Differentiating sepsis from uncomplicated infection is important to ensure rapid treatment for those who require it, balanced against the harms of overtreatment for those who do not. As such assessing accuracy of existing tools is a potentially valuable exercise. However, there are some methodological problems with this study that need considering Usually a diagnostic study aims to assess a new or index test against a reference/gold standard test against the full breadth clinical presentations assessing accuracy of a particular outcome (diagnosis, ICU admission, mortality). 1. The first problem with this study is that only patients could enter if they met SIRS criteria. This skews any findings on diagnostic accuracy as patients with infection without SIRS (older patients, patients on beta-blockers, atupical presentations are excluded – they usually have poorer outcomes). To truly comment on diagnostic accuracy of SIRS, you would need to include all patients presenting to ED with clinician suspected infection 2. A second problem with this study is that the two tools were made for different reasons and as such a diagnostic study is problematic. Sepsis-2 placed substantial weight upon clinical judgment of an infectious aetiology AND the established SIRS components. As outlined by the authors SIRS attracted criticism for lack of specificity. As such SIRS has been more often used as a screening test (more sensitive than specific). The Sepsis-3 criteria now define sepsis as an infection-related acute increase of 2 or more points from baseline in the Sequential Organ Failure Assessment (SOFA) score. This is difficult to assess in the ED and as such the 'quick-SOFA' score (qSOFA) was suggested for use at the bedside. Importantly, when derived its predictive value was not specifically assessed in ED patients, and sensitivity for detecting adverse outcomes outside the ICU has been found to be limited. As such it was never recommended for this purpose. Rather qSOFA should be considered a severity assessment tool or prognostic tool to identify those with infection at particularly high risk of deterioration. 3. A third problem is the outcome definition. Having 'sepsis' was based on culture. Although reproducible and well-intended, most diagnostic studies on sepsis have different definitions. On top of this sputum and throat cultures (colonisation rather than infection can be found(are not evidence-based and a substantial proportions of sepsis can be culture negative.

	4. The adherence to (two different?) treatment protocols is problematic. There is indeed evidence that bundles of care are associated with better outcomes, although it is very unclear what parts of the bundle are responsible for this. It may well be that the fact the patient receives closer observation when a sepsis pathway is started, rather than benefit of individual bundle elements. The authors may be well aware of significant criticism on mandated sepsis bundles (mainly in the USA) 5. Lastly, the division in the 5 groups is well intended, but of uncertain utility. In essence the authors are looking at 2 tools intended and designed for a different purpose (screening vs prognostication) and the adherence to the protocol would only be of relevance of the outcome of interest was a patient centered outcome such as ICU admission or mortality (as the protocol adherence has no bearing on the diagnostic accuracy of sepsis) In summary – The patient selection requiring SIRS to be +ve and comparing diagnostic accuracy of two tools that were not intended/derived for the population applied to as well as the definition of sepsis make it very hard to interpret the diagnostic accuracy data. As it is very hard for the reader to assess what this data means, it would be better to be written up as a purely descriptive study. Please see the references to these 2 recent review papers which may inform any revision or new submission Shetty et al Emerg Med Australas. 2018 Feb;30(1):4-12. doi: 10.1111/1742-6723.12924. Epub 2018 Jan 16. Williams et al. Emerg Med Australas. 2017 Dec;29(6):619-625. doi: 10.1111/1742-6723.12886. Epub 2017 Nov 2.

VERSION 1 – AUTHOR RESPONSE

Major points:

1. As suggested by reviewer 1, we redefined the outcome as identifying culture-positive sepsis.
2. We adapted the aim of our study, to be a more descriptive study. We intended to give a representation of the daily practice of two EDs in the Netherlands, where SIRS criteria and qSOFA score were used differently as screening tool for culture-positive sepsis. Although qSOFA was initially not developed as a screening instrument but as a risk stratification for sepsis prognosis, it is used as such in many Dutch EDs (most of the Dutch EDs use qSOFA in combination as both a screening instrument and as risk stratification). Because qSOFA is used as such, it is important to know its value as a screening tool.
3. There is some confusion about the treatment protocols, which we have defined more clearly. Both hospitals used the same treatment protocol, the worldwide Surviving Sepsis Campaign protocol (which consist of antibiotics, fluid resuscitation and oxygen supplementation). Only different criteria were used for the start of this protocol (SIRS criteria in Maxima Medical Centre and qSOFA score in Amphia Hospital).
4. According to reviewer 2: it is a serious limitation of our study to not include patients suspected for infection with negative SIRS criteria and qSOFA score. Because of this we could only calculate the positive predictive value of both screening instruments. We described this limitation more clearly in the discussion part. Unfortunately, it is not possible for us to re-include the patients with negative

SIRS criteria and qSOFA score.
5. Both reviewers indicated that patient groups 4 and 5 are too small. The small number of these groups is because it is unusual for patients to be SIRS negative and qSOFA positive given the type of vital parameters contained in these screening instruments. This means that with a larger study population, these two groups will not grow in the same line. Due to the small number of these groups, we agreed that it is not possible to make a statement about the results and redefined this in our study.

Minor points:
1. We added the approval of the Medical Research Ethics Committees Máxima Medical Centre in our article.
2. We added the use of the STARD checklist when writing our report.

We hope that our modifications render our manuscript in its current form suitable for publication in BMJ Open.

VERSION 2 – REVIEW

REVIEWER	Shi, Haolun The University of Hong Kong
REVIEW RETURNED	26-Feb-2021

GENERAL COMMENTS	No comment

REVIEWER	Athens, Josie University of Otago, Preventive and Social Medicine
REVIEW RETURNED	03-Mar-2021

GENERAL COMMENTS	I did not know I was making a review of a second submission till I found the corrected version in the second half of the document. Reading the first half, for me, I found of concern to not see anything about Ethics. I am very happy that was addressed in the second version. The manuscript is very clear, the statistical analysis is straightforward and the results were well discussed. I appreciate that the authors acknowledged their limitations and were careful when deriving conclusions. The only observation I have is about Figure 3. The quality of what I saw was poor and in no way appropriate for publication. Furthermore, there is no need to start the y-axis at zero, but that is optional, just make sure that the final figure if of high quality.

REVIEWER	Der-Martirosian, Claudia Veterans Emergency Management Evaluation Center
REVIEW RETURNED	04-Mar-2021

GENERAL COMMENTS	Statistical section: Thank you for the opportunity to review this manuscript. The authors write: "Differences in baseline characteristics between hospitals and study groups were tested using Chi-square statistics and Mann-Whitney U tests for categorical respectively continuous variables." The phrase "categorical respectively continuous variables" does not make sense. I think the authors meant to say: "categorical and continuous variables, respectively."

	"p-value of <0.05" should replace "p-value of <0,05" in the statistical analysis section. The word "respectively" is misused (or misplaced in the sentences) throughout the manuscript. Please edit carefully.

VERSION 2 – AUTHOR RESPONSE

The manuscript has been improved according to the suggestions of the reviewers:

Minor points:

1. We have improved the quality of Figure 3 compared to its previous version.
2. We replaced ‘p-value of <0,05’ with ‘p-value of <0.05’ in the statistical analysis section.
3. We have reconsidered and corrected the sentences with the word ‘respectively’ and used it carefully throughout the manuscript.

We hope that our modifications render our manuscript in its current form suitable for publication in BMJ Open.