BMJ Open

Retrospective study on the possible existence of a treatment paradox in sepsis scores in the emergency department

Jan Willem Uffen ![ORCID],[1] Harriet van Goor,[1] Johannes Reitsma,[2]
Jan Jelrik Oosterheert,[3] Marieke de Regt,[4] Karin Kaasjager[1]

ABSTRACT

Objective The quick Sequential Organ Failure Assessment (qSOFA) is developed as a tool to identify patients with infection with increased risk of dying from sepsis in non-intensive care unit settings, like the emergency department (ED). An abnormal score may trigger the initiation of appropriate therapy to reduce that risk. This study assesses the risk of a treatment paradox: the effect of a strong predictor for mortality will be reduced if that predictor also acts as a trigger for initiating treatment to prevent mortality.

Design Retrospective analysis on data from a large observational cohort.

Setting ED of a tertiary medical centre in the Netherlands.

Participants 3178 consecutive patients with suspected infection.

Primary outcome To evaluate the existence of a treatment paradox by determining the influence of baseline qSOFA on treatment decisions within the first 24 hours after admission.

Results 226 (7.1%) had a qSOFA ≥2, of which 51 (22.6%) died within 30 days. Area under receiver operating characteristics of qSOFA for 30-day mortality was 0.68 (95% CI 0.61 to 0.75). Patients with a qSOFA ≥2 had higher odds of receiving any form of intensive therapy (OR 11.4 (95% CI 7.5 to 17.1)), such as aggressive fluid resuscitation (OR 8.8 95% CI 6.6 to 11.8), fast antibiotic administration (OR 8.5, 95% CI 5.7 to 12.3) or vasopressic therapy (OR 17.3, 95% CI 11.2 to 26.8), compared with patients with qSOFA <2.

Conclusion In ED patients with suspected infection, a qSOFA ≥2 was associated with more intensive treatment. This could lead to inadequate prediction of 30-day mortality due to the presence of a treatment paradox.

Trial registration number 6916.

INTRODUCTION

In absence of a gold diagnostic standard for sepsis, the Third International Consensus Definition Task Force (Sepsis-3) introduced the quick Sequential Organ Failure Assessment (qSOFA) as a prognostic tool for sepsis outside the intensive care unit (ICU).[1] This score was developed and validated in

Strengths and limitations of this study

► This study suggests the existence of a treatment paradox in the interpretation of the quick Sequential Organ Failure Assessment in sepsis care.
► This study addresses the consequences of a treatment paradox when physicians or researchers are not aware of this phenomenon.
► The concept of a treatment paradox in the field of sepsis remains a suggestion and cannot be proven.
► Results of this study could have been influenced by delayed administration in the emergency department, leading to a possible underestimation of the effect of abnormal parameters on antbiotic administration.

large retrospective cohorts of patients with suspected infection to identify patients with an increased risk of dying.[2–8] Based on the fundamental idea that sepsis is more severe and has a higher mortality rate than ordinary infections, this prognostic score was introduced as a proxy for a diagnostic tool to enhance sepsis recognition and guide treatment decisions.[1]

In clinical medicine, prognostic scores are often used to identify patients at a certain risk of disease or unwanted outcome who might benefit from particular interventions. Ideally, these prognostic scores are developed using patient populations that have not been treated for the outcome of interest.[9] However, this is often impossible due to ethical considerations. Furthermore, if the risk of disease or unwanted outcome is high, additional treatments are often initiated. This introduces the risk of a treatment paradox: when a strong prognostic factor of an adverse outcome triggers an effective treatment, the incidence of this outcome will be reduced.[10 11] In this situation, the prognostic factor that initiated the treatment will appear to have a poorer
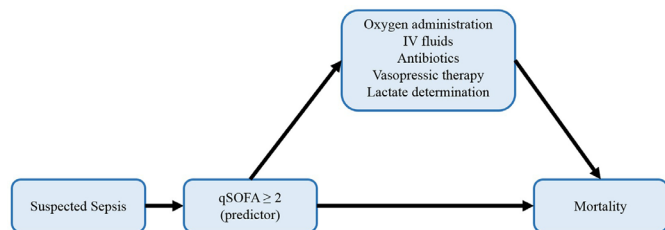
Figure 1 Graphical representation of a possible treatment paradox when using qSOFA as a clinical score in sepsis care. The observed predictor outcome relation becomes a combination of the direct effect of a qSOFA ≥2 and of the indirect effect of oxygen administration, intravenous fluid therapy, antibiotic therapy, vasopressic therapy and lactate determination. qSOFA, quick Sequential Organ Failure Assessment.

prognostic performance than it actually has (figure 1). Two factors are essential for a treatment paradox to occur: (1) the prognostic factor has a strong relationship with the outcome and (2) when the prognostic factor is present, it triggers an effective treatment. The presence of this phenomenon may lead to a biased underestimation of adverse outcomes when prognostic scores, developed in a treated population, are applied to treatment-naive patients. In other words, patients with a high benefit of treatment may not be recognised properly, posing the risk of under treatment. The treatment paradox has been recognised in several medical fields, such as obstetrics and cardiovascular management.[11–14]

With the introduction of prognostic sepsis scores, the same problem may have been introduced in the field of sepsis. The qSOFA contains baseline characteristics that are likely to alarm the treating physician to act when abnormal, such as a high respiratory rate and a low systolic blood pressure.[1] Moreover, qSOFA has been developed and validated in large retrospective cohorts in which patients with infection and sepsis were treated, thereby introducing the risk of a treatment paradox.

In this study, we explore the potential existence of a treatment paradox in sepsis care by analysing the two essential factors required: (1) the relationship between the baseline qSOFA in the ED and 30-day mortality and (2) the relation between the qSOFA and subsequent treatment decisions within the first 24 hours after admission.

METHODS
Study design and setting
Analysis was performed on data from the prospective SePsis in the ACutely ill patients in the Emergency department (SPACE) cohort.[15] This cohort consists of all consecutive patients that presented to the ED of the University Medical Center Utrecht for internal medicine with suspected infection between 13 September 2016 and 13 September 2018 and started shortly after the publication of Sepsis-3. This ED has an annual presentation number of 23 000 cases. Patient data collected in this cohort consist of clinical data on presentation in the ED

and, if applicable, hospital admission. Furthermore, data is collected on diagnostic tests, treatment and follow-up up to 30 days. The SPACE cohort was registered in the Netherlands Trial Register.

Population and data collection
The SPACE cohort consists of all consecutive patients who meet the following criteria: (1) ≥18 years or older; (2) presentation at the ED with suspected infection defined by the treating physician in the ED; and (3) registration in the ED for the internal medicine department or its subspecialties: oncology, rheumatology, immunology, haematology, nephrology, endocrinology, gastroenterology, infectious disease and vascular medicine. All patients received standard care.

All patients in the ED were treated according to a local sepsis clinical pathway. This pathway includes qSOFA and the Systemic Inflammatory Response Syndrome (SIRS) criteria. Patients suspected of sepsis by either positive SIRS, qSOFA or clinical suspicion receive care following different diagnostic and treatment steps based on international guidelines. For example, this includes resuscitation according to ABCDE method, diagnostic tests including blood culture withdrawal and administration of empiric antibiotics <1 hour. It also advices the use of vasopressor when a patient is not responding to fluid resuscitation of 1.5 L of intravenous fluid.

The qSOFA was automatically calculated and reported in the electronic health record (EHR) system after the treating physician answered a non-obligatory question if he or she clinically suspected infection or sepsis. The calculation of the qSOFA was based on the first available recorded data and was considered positive in case of a qSOFA ≥2. Independent-trained physicians analysed all EHRs on documented suspected infection or sepsis. If the infection and sepsis questions were not answered by the treating physician, the independent physician marked the questions positive when respectively infection or sepsis was recorded by the ED physician as (differential) diagnosis in the ED patient record.

General patient information, data on hospital or ICU or medium care (MC) admission, vital signs, laboratory testing and mortality were automatically extracted from the EHR.

Data on comorbidities (categorised using the Charlson Comorbidity Index (CCI)),[16] immunocompromised status and information on treatments administrated in the ED (administration of intravenous fluids and antibiotics, time to first antibiotics, oxygen therapy and use of vasopressor agents) were manually extracted from the EHR by researchers, using a predefined set of well-described definitions. If Glasgow coma scale was not registered, free text notes by the treating ED physician on the mental status were used. The (differential) diagnosis at admission and diagnosis at discharge were retrieved from the ED record and hospital discharge letter, respectively. These diagnoses were reviewed on correctness and accuracy by a standardised independent review of the medical

record by the principal investigators, using predefined definitions. This review was based on symptoms, vital signs, laboratory results, radiology results and microbiology results and was standardised for most common infections.

Outcome measurements

The potential presence of a treatment paradox was investigated by: (1) evaluating the prognostic accuracy of the qSOFA for 30-day mortality within this cohort and by (2) analysing the relationship between a positive qSOFA and abnormal vital parameters imbedded in the qSOFA and the intensity of initiated therapy within the first 24 hours after ED presentation. Five therapy elements were investigated and were considered intensive (vs less intensive) in the following situations: (1) volume of intravenous fluid resuscitation within the first 3 hours on ED admission, (2) oxygen administration in the ED, (3) the use of vasopressors within the first 24 hours on admission, (4) a lactate measurement in the ED (which is suggested by the clinical pathway when at least one qSOFA parameter is abnormal) and (5) antibiotic treatment <1 hour. The latter was determined by calculating the time between arrival in the ED and the first antibiotic administration, registered by the nurse in the EHR. These interventions were used as a proxy for intensive therapy. Choice of these interventions do not suggest these are by means beneficial to this group of patients.

Patient and public involvement

Patients and/or public were not involved in this research.

Statistical analysis

For prognostic validation of the qSOFA for 30-day mortality a sensitivity/specificity for a qSOFA ≥2 and an area under receiver operating characteristics (AUROC) curve for qSOFA were calculated.

Binary logistic regression analyses were performed to study the relationship between qSOFA as binary covariate and the choice of a positive qSOFA on initiated intensive therapy as outcome. A separate regression analysis was done per specific therapy element and for a combined outcome defined as receiving at least one form of intensive therapy, resulting in six analyses. Age and CCI were added as potential confounders to all models. Relationships were expressed as OR with 95% CI, and p values were derived. A p value <0.05 was considered as statistically significant. Multicollinearity of determinants was explored by deriving Spearman's correlation coefficients. Hosmer-Lemeshow goodness of fit test was used to determine the fit of the extracted model.

RESULTS

In total, 3178 consecutive patients were included. Patient characteristics are described in table 1. In this cohort, the most common infection diagnosed in the ED was a lower respiratory tract infection (20.7%, n=658). In 338 (10.6%)

patients, an alternative non-infectious diagnosis was made during hospital admission or outpatient follow-up, most commonly a side effect of medication (24.0%, n=81). These patients were included in the analysis, because they were treated as suspected infectious in the ED.

Of all patients, 1089 (34.3%) were immunocompromised. Two-thirds (n=2134, 67.1%) of the patients were admitted to the hospital, and in 2174 (68.4%) patients, antibiotics were started in the ED. In total, 315 (9.9%) patients were admitted to the ICU or MC at any point during admission and 195 (6.1%) patients died within 30 days.

Prognostic accuracy of qSOFA for 30-day mortality

The risk of dying within 30 days after ED presentation increased from 6.1% (195/3178) in all patients to 22.6% (51/226) in the subgroup of patients with a baseline qSOFA ≥2. For 30-day mortality, a qSOFA ≥2 had a sensitivity of 0.26 (95% CI 0.20 to 0.33) and specificity of 0.94 (95% CI 0.93 to 0.95). The qSOFA had an AUROC curve of 0.68 (95% CI 0.61 to 0.75).

Treatment paradox

Patients with a baseline qSOFA ≥2 or with an abnormal individual element of qSOFA more frequently received any form of intensive therapy (table 2). Table 3 shows that a qSOFA ≥2 was independently associated with more frequently: (1) receiving antibiotics within 1 hour (OR 8.5 (95% CI 5.7 to 12.3)), (2) receiving more than 1 L intravenous fluids (OR 8.8 (95% CI 6.6 to 11.8)), (3) receiving vasopressor therapy (OR 17.3 (95% CI 11.2 to 26.8)), (4) receiving oxygen therapy (OR 6.4 (95% CI 4.7 to 8.7)) and (5) lactate measurement in the ED (OR 6.9 (95% CI 5.0 to 9.4)) compared with patients who had a qSOFA <2. Furthermore, there was an increased odds of 11.4 (95% CI 7.5 to 17.1) for patients with a qSOFA ≥2 of receiving at least one form of intensive therapy compared with patients with a qSOFA <2. Table 3 also shows that abnormal individual elements of the qSOFA were associated with more forms of intensive therapy.

DISCUSSION

With this article, we under scribe the theoretical existence of a treatment paradox in sepsis care in the ED by demonstrating that an abnormal qSOFA and abnormal individual elements of the qSOFA are associated with intensive treatment in the ED in patients with suspected infection. The problem lies in the fact that the prognostic qSOFA score was developed and validated in a cohort of patients that received treatment on clinical indication. Assuming these therapeutic interventions were, at least partly, effective in treating the suspected infection and sepsis, the incidence of adverse outcomes will have been reduced. As a consequence, the effect of strong predictors of mortality that trigger effective treatment will be underestimated. Briefly, the qSOFA is especially suited to identify patients that die despite treatment. This

Table 1 Patients characteristics of all patients and patients with a qSOFA ≥2

	All patients	qSOFA <2	qSOFA ≥2
	n=3178	n=2952	n=226
Age (IQR)	61 (48–70)	61 (46–70)	66 (58–75)
Gender (% male)	51.5	52.3	57.1
Infectious diagnosis on presentation, n (%)			
LRTI	658 (20.7)	565 (19.1)	93 (41.2)
Urinary tract infection	582 (18.3)	536 (18.2)	46 (20.4)
Viral upper airway infection	535 (16.8)	523 (17.7)	12 (5.3)
GI infection	475 (14.9)	459 (15.5)	16 (7.1)
Skin infection	232 (7.3)	217 (7.4)	15 (6.6)
Immunocompromised, n (%)	1089 (34.3)	1022 (34.6)	67 (29.6)
Charlson comorbidity index (median, IQR)	4 (2–7)	4 (2–7)	5 (4–8)
Admission to hospital, n (%)	2134 (67.1)	1916 (64.9)	218 (96.5)
Antibiotics administrated, n (%)			
Inpatient	1618 (50.9)	1431 (48.5)	187 (82.7)
Outpatient	556 (17.5)	545 (18.5)	11 (4.9)
No antibiotics	1004 (31.6)	976 (33.1)	28 (12.4)
Non-infectious diagnosis on discharge, n (%)	338 (10.6)	324 (11.0)	14 (6.2)
Medication side effect	81 (24.0)	78 (24.1)	3 (21.4)
Malignancy	56 (16.6)	52 (16.0)	4 (28.5)
Autoimmune disease	30 (8.9)	29 (9.0)	1 (7.1)
ICU/MCU admission, n (%)	315 (9.9)	215 (7.3)	100 (44.2)
Died within 30 days, n (%)	195 (6.1)	144 (4.9)	51 (22.6)

GI, gastrointestinal; ICU, intensive care unit; LRTI, lower respiratory tract infection; MCU, medium care unit; qSOFA, quick Sequential Organ Failure Assessment.

could potentially lead to an underestimation of adverse outcomes in treatment-naïve patients that actually benefit from treatment the most.

An illustrative example of the risks of ignoring the treatment paradox in prognostic models can be found in the field of obstetrics.[17] A retrospective study aimed to develop a prediction model for adverse maternal outcomes in suspected pre-eclampsia failed to identify maternal hypertension as a risk factor for adverse outcome due to a treatment paradox.[13] In the study cohort, maternal hypertension was such a strong trigger for physicians to start an effective treatment that significantly less adverse events occurred. As a consequence, the statistical inference between maternal hypertension and adverse outcomes completely disappeared, and this well-known risk factor was not included in the prognostic model. However, ignoring a strong risk factor such as maternal hypertension in pre-eclampsia in new treatment-naïve patients would certainly lead to undertreatment and adverse outcomes. Although the treatment paradox effect in sepsis will probably be less strong than in pre-eclampsia, because sepsis is a far more heterogeneous syndrome with more heterogeneity in treatment effects, the results of our study support the presence of the effect.

In the constant search for new screening tools in the field of sepsis, studies developing new or validating existing prognostic models rarely address the possible existence of a treatment paradox in sepsis recognition and treatment. Treatment paradoxes in the field of sepsis are likely to occur in different types of prognostic scores. Two meta-analyses on the qSOFA, Early Warning Scores and the SIRS criteria, only briefly recognise and address the risk of bias introduced by the treatment paradox in all studies included in their analysis, without discussing the further consequences. None of the individual studies in the meta-analyses discuss the possible existence of a treatment paradox.[18 19] One retrospective analysis aimed to improve the National Early Warning Score by adding inflammatory blood marker addresses that their results could have been influenced by a treatment paradox, but this study was not specifically performed in patients at risk for sepsis.[20] Future research on validating existing scores and developing new prediction models in sepsis should address these methodological issues by using other data sources like additional testing, follow-up, response to treatment or expert panels estimating sepsis risk and combining the imperfect information through latent class models and/or measurement error models. A

Table 2 Therapy aggressiveness in patients with a qSOFA ≥2 and abnormal elements of the qSOFA and mortality rates per different therapy element

	Total (n=3178)	Antibiotics >1hour or no antibiotics (n=3022)	Antibiotics <1hour (n=156)	Fluids <1L (n=2693)	Fluids ≥1L (n=485)	No vasopressor therapy (n=3084)	Vasopressor therapy (n=94)	No oxygen therapy (n=2812)	Oxygen therapy (n=366)	No lactate determined* (n=2219)	Lactate determined* (n=959)
qSOFA ≥2	7.1 (226)	7.0 (142)	35.9 (56)	3.7 (100)	26.0 (126)	5.8 (178)	51.1 (48)	4.6 (129)	26.5 (97)	2.7 (60)	17.3 (166)
Respiratory rate ≥22/min	19.7 (626)	18 (543)	53.2 (83)	16.0 (432)	40.0 (194)	18.4 (566)	63.8 (60)	15.3 (429)	53.8 (197)	11.0 (243)	39.9 (383)
SBP ≤100mm Hg	10.1 (320)	9.0 (271)	31.4 (49)	6.0 (162)	32.6 (158)	8.9 (273)	50.0 (47)	8.3 (234)	23.5 (86)	6.9 (153)	17.4 (167)
GCS <15	8.1 (256)	6.9 (210)	29.5 (46)	6.1 (164)	19.0 (92)	7.2 (221)	37.2 (35)	5.9 (166)	24.6 (366)	3.9 (86)	17.7 (170)
30-day mortality	6.1 (195)	5.8 (177)	11.5 (18)	4.8 (130)	13.4 (65)	5.5 (171)	25.5 (24)	4.7 (132)	17.2 (63)	4.2 (93)	10.6 (102)

Percentage of total (number).
*Lactate determination in the emergency department.
GCS, Glasgow coma scale; qSOFA, quick Sequential Organ Failure Assessment; SBP, systolic blood pressure.

Table 3 Showing the association between a qSOFA ≥2 and the choice for intensive therapy

	Outcome	Antibiotics<1hour.	≥1L intravenous fluids	Vasopressor therapy	Oxygen therapy	Lactate measurement*	Combined outcome†
Covariates	qSOFA ≥2‡	8.5 (5.7 to 12.3)	8.8 (6.6 to 11.8)	17.3 (11.2 to 26.8)	6.4 (4.7 to 8.7)	6.9 (5.0 to 9.4)	11.4 (7.5 to 17.1)
	Respiratory Rate ≥22/min§	3.5 (2.5 to 5.0)	2.4 (1.9 to 3.0)	4.1 (2.5 to 6.6)	3.6 (2.8 to 4.8)	3.3 (2.7 to 4.1)	4.0 (3.3 to 5.0)
	SBP ≤100mm Hg¶	3.2 (2.2 to 4.7)	5.9 (4.4 to 7.8)	6.8 (4.2 to 10.8)	2.3 (1.7 to 3.2)	1.9 (1.5 to 2.5)	2.9 (2.2 to 3.9)
	GCS <15**	2.9 (1.9 to 4.4)	2.1 (1.5 to 2.9)	3.6 (2.2 to 6.0)	2.3 (1.7 to 3.2)	3.0 (2.2 to 4.1)	3.2 (2.3 to 4.5)

Results from six different backward stepwise logistic regression analyses. Correction was made for Charlson Comorbidity Index and age. Associations are presented in ORs and 95% CI.
*Lactate measurement in the emergency department.
†Combined outcome: at least one form of intensive therapy.
‡Correction made for: age and comorbidities.
§Correction made for: age, comorbidities, SBP and GCS.
¶Correction made for: age, comorbidities, respiratory rate and GCS.
**Correction made for: age, comorbidities, respiratory rate and SBP.
GCS, Glasgow coma scale; qSOFA, quick Sequential Organ Failure Assessment; SBP, systolic blood pressure.

recent framework by van Geloven et al[21] is a useful starting point to evaluate the various types of predictions that can be made that explicitly incorporate the use of treatments.

This study has several strengths. First, we address an important epidemiological phenomenon in using the qSOFA in sepsis care and that probably applies for all scoring systems used to predict patient outcomes that can be prevented by initiating effective treatment and support this by illustrating this with real patient data. Second, we discuss the consequences of a treatment paradox when clinicians or researchers are not aware of this phenomenon. Furthermore, the study population exists of a heterogeneous group of patients most at risk for developing sepsis. Therefore, results obtained from this study are applicable in daily practice.

This study has several limitations. The true existence of a treatment paradox cannot be proven. We only provided evidence for the requirements for a potential treatment paradox in ED sepsis care by showing the prognostic accuracy of qSOFA and 30-day mortality and the relation between qSOFA and the initiation of intensive treatment decisions. Furthermore, only 35.9% of the patients who had a qSOFA ≥2 received antibiotic treatment within 1 hour, despite antibiotic treatment within 1 hour is mandatory according to local protocols. This is probably due to a delay in administrating antibiotics to the patients and registration of the antibiotic in the EHR. The actual antibiotic administration could have happened earlier than registration times. This probably led to an underestimation of the association between a qSOFA ≥2 and the administration of antibiotics within 1 hour after presentation in the ED.

Lastly, the SPACE cohort started shortly after the introduction of Sepsis-3. This could have resulted in unfamiliarity with the qSOFA during early stages of the study.

The qSOFA should be used as intended. It is a prognostic score for predicting short term (in-hospital) mortality in patients with suspected of confirmed infection in non-ICU settings. It should not be used as a diagnostic tool for the presence or absence of sepsis. Many studies have validated the qSOFA on their own patient cohorts, resulting in prognostic accuracy measures comparable with our findings. The question arises if sepsis recognition should rely on the use of a prognostic score with many intrinsic problems. Other clinical scores, like early warning scores, have been suggested of use in sepsis care.[22 23] Future studies on these scores, aimed at external validation and/ or updating, should explicitly indicate how treatments will be incorporated in these models and what kind of predictions will be made by the model. Other modelling approaches and other data sources are indispensable to build better models for predicting sepsis or mortality. Until then, clinicians should be aware of how a treatment paradox affects the interpretation of the qSOFA in sepsis care. Keeping this in mind, combined with careful consideration of its results within the complex of clinical data and clinical bedside judgement, the qSOFA (or any other clinical score) might still be helpful in recognising and treating sepsis.

Author affiliations
[1]Department of Internal Medicine and Acute Medicine, University Medical Centre Utrecht, Utrecht, The Netherlands
[2]Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[3]Department of Internal Medicine and Infectious Diseases, University Medical Center Utrecht, Utrecht, The Netherlands
[4]Department of Internal Medicine, Onze Lieve Vrouwe Gasthuis, Amsterdam, Noord-Holland, The Netherlands

ORCID iD
Jan Willem Uffen http://orcid.org/0000-0001-7196-1229

REFERENCES
1 Singer M, Deutschman CS, Seymour CW, et al. The third International consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016;315:801–10.
2 Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of clinical criteria for sepsis: for the third International consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;315:762–74.
3 Askim Åsa, Moser F, Gustad LT, et al. Poor performance of quick-SOFA (qSOFA) score in predicting severe sepsis and mortality - a prospective study of patients admitted with infection to the emergency department. *Scand J Trauma Resusc Emerg Med* 2017;25:56.
4 Williams JM, Greenslade JH, McKenzie JV, et al. Systemic inflammatory response syndrome, quick sequential organ function assessment, and organ dysfunction: insights from a prospective database of ED patients with infection. *Chest* 2017;151:586–96.

5 González Del Castillo J, Julian-Jiménez A, González-Martínez F, *et al*. Prognostic accuracy of SIRS criteria, qSOFA score and GYM score for 30-day-mortality in older non-severely dependent infected patients attended in the emergency department. *Eur J Clin Microbiol Infect Dis* 2017;36:2361–9.

6 Goulden R, Hoyle M-C, Monis J, *et al*. qSOFA, SIRS and news for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. *Emerg Med J* 2018;35:345–9.

7 Moskowitz A, Patel PV, Grossestreuer AV, *et al*. Quick sequential organ failure assessment and systemic inflammatory response syndrome criteria as predictors of critical care intervention among patients with suspected infection. *Crit Care Med* 2017;45:1813–9.

8 Freund Y, Lemachatti N, Krastinova E, *et al*. Prognostic accuracy of sepsis-3 criteria for in-hospital mortality among patients with suspected infection presenting to the emergency department. *JAMA* 2017;317:301–8.

9 Moons KGM, Royston P, Vergouwe Y, *et al*. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375–20.

10 Schuit E, Groenwold RHH, Harrell FE, *et al*. Unexpected predictor-outcome associations in clinical prediction research: causes and solutions. *Can Med Assoc J* 2013;185:E499–505.

11 Pajouheshnia R, Peelen LM, Moons KGM, *et al*. Accounting for treatment use when validating a prognostic model: a simulation study. *BMC Med Res Methodol* 2017;17:103.

12 Westerhuis MEMH, Schuit E, Kwee A, *et al*. Prediction of neonatal metabolic acidosis in women with a singleton term pregnancy in cephalic presentation. *Am J Perinatol* 2012;29:167–74.

13 Cheong-See F, Allotey J, Marlin N, *et al*. Prediction models in obstetrics: understanding the treatment paradox and potential solutions to the threat it poses. *BJOG* 2016;123:1060–4.

14 Roffi M, Mukherjee D. Treatment-risk paradox in acute coronary syndromes. *Eur Heart J* 2018;39:3807–9.

15 Uffen JW, Oomen P, de Regt M, *et al*. The prognostic value of red blood cell distribution width in patients with suspected infection in the emergency department. *BMC Emerg Med* 2019;19:76.

16 Charlson ME, Pompei P, Ales KL, *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.

17 von Dadelszen P, Payne B, Li J, *et al*. Prediction of adverse maternal outcomes in pre-eclampsia: development and validation of the fullPIERS model. *Lancet* 2011;377:219–27.

18 Maitra S, Som A, Bhattacharjee S. Accuracy of quick sequential organ failure assessment (qSOFA) score and systemic inflammatory response syndrome (SIRS) criteria for predicting mortality in hospitalized patients with suspected infection: a meta-analysis of observational studies. *Clin Microbiol Infect* 2018;24:1123–9.

19 Hamilton F, Arnold D, Baird A, *et al*. Early warning scores do not accurately predict mortality in sepsis: a meta-analysis and systematic review of the literature. *J Infect* 2018;76:241–8.

20 Eckart A, Hauser SI, Kutz A, *et al*. Combination of the National early warning score (news) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study. *BMJ Open* 2019;9:e024636.

21 van Geloven N, Swanson SA, Ramspek CL, *et al*. Prediction meets causal inference: the role of treatment in clinical prediction models. *Eur J Epidemiol* 2020;35:619–30.

22 Churpek MM, Snyder A, Han X, *et al*. Quick sepsis-related organ failure assessment, systemic inflammatory response syndrome, and early warning scores for detecting clinical deterioration in infected patients outside the intensive care unit. *Am J Respir Crit Care Med* 2017;195:906–11.

23 de Groot B, Stolwijk F, Warmerdam M, *et al*. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. *Scand J Trauma Resusc Emerg Med* 2017;25:91.