ORIGINAL ARTICLE

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

Matthew M. Churpek[1,2], Ashley Snyder[1], Xuan Han[1], Sarah Sokol[3], Natasha Pettit[3], Michael D. Howell[1,2], and Dana P. Edelson[1,2]

[1]Department of Medicine, [2]Center for Healthcare Delivery Science and Innovation, and [3]Department of Pharmacy, University of Chicago, Chicago, Illinois

Abstract

Rationale: The 2016 definitions of sepsis included the quick Sepsis-related Organ Failure Assessment (qSOFA) score to identify high-risk patients outside the intensive care unit (ICU).

Objectives: We sought to compare qSOFA with other commonly used early warning scores.

Methods: All admitted patients who first met the criteria for suspicion of infection in the emergency department (ED) or hospital wards from November 2008 until January 2016 were included. The qSOFA, Systemic Inflammatory Response Syndrome (SIRS), Modified Early Warning Score (MEWS), and the National Early Warning Score (NEWS) were compared for predicting death and ICU transfer.

Measurements and Main Results: Of the 30,677 included patients, 1,649 (5.4%) died and 7,385 (24%) experienced the composite outcome (death or ICU transfer). Sixty percent (n = 18,523) first met the suspicion criteria in the ED. Discrimination for in-hospital mortality was highest for NEWS (area under the curve [AUC], 0.77; 95% confidence interval [CI], 0.76–0.79), followed by MEWS (AUC, 0.73; 95% CI, 0.71–0.74), qSOFA (AUC, 0.69; 95% CI, 0.67–0.70), and SIRS (AUC, 0.65; 95% CI, 0.63–0.66) ($P < 0.01$ for all pairwise comparisons). Using the highest non-ICU score of patients, ≥2 SIRS had a sensitivity of 91% and specificity of 13% for the composite outcome compared with 54% and 67% for qSOFA ≥2, 59% and 70% for MEWS ≥5, and 67% and 66% for NEWS ≥8, respectively. Most patients met ≥2 SIRS criteria 17 hours before the combined outcome compared with 5 hours for ≥2 and 17 hours for ≥1 qSOFA criteria.

Conclusions: Commonly used early warning scores are more accurate than the qSOFA score for predicting death and ICU transfer in non-ICU patients. These results suggest that the qSOFA score should not replace general early warning scores when risk-stratifying patients with suspected infection.

Keywords: systemic inflammatory response syndrome; sepsis; organ dysfunction scores; early warning scores; qSOFA

At a Glance Commentary

Scientific Knowledge on the Subject: The quick Sepsis-related Organ Failure Assessment (qSOFA) score is a bedside tool that was recommended for use by the recent Third International Consensus Definitions Task Force to identify high-risk patients outside of the intensive care unit (ICU). qSOFA was found to be more accurate than the Systemic Inflammatory Response Syndrome (SIRS) criteria for predicting mortality and ICU transfer in patients outside the ICU. However, the qSOFA score has yet to be validated outside of the original publication and has not been compared with early warning scores already in widespread use.

What This Study Adds to the Field: We found that commonly used early warning scores are more accurate than the qSOFA score for predicting in-hospital mortality and ICU transfer among patients with suspected infection outside the ICU. These results suggest that the qSOFA score should not replace previously developed early warning scores already in use across the United States and Europe.

Sepsis contributes to up to one-half of hospital deaths and is associated with more than $24 billion in costs in the United States annually (1, 2). Previous consensus definitions included the Systemic Inflammatory Response Syndrome (SIRS) criteria, which consisted of temperature, white blood cell count, heart rate, and respiratory rate (3, 4). However, SIRS criteria have been criticized for their poor specificity, with 90% of intensive care unit (ICU) patients and 50% of general ward patients meeting the criteria at some point during their hospitalization (5–7). This led the authors of the 2016 sepsis guidelines to use a more data-driven approach to developing the definition of sepsis (8, 9).

Because most cases of sepsis present in the emergency department (ED) and on the wards rather than the ICU (10), the 2016 guidelines included a new tool that was derived specifically to prompt clinicians to consider possible sepsis (9). This model, called the quick Sepsis-related Organ Failure Assessment (qSOFA), requires derangements in systolic blood pressure, mental status, and respiratory rate, and was found to be more accurate than SIRS for predicting adverse events (9). Although supported by analyses of large data sets, the new definitions have been criticized because they identify patients too late in their course of illness (11). Furthermore, qSOFA was not compared with risk stratification tools that are already commonly implemented in clinical practice outside the ICU, such as the Modified Early Warning Score (MEWS) (12) and the National Early Warning Score (NEWS) (13). There are substantial clinical and operational benefits to using scores that are already in place if they have similar performance characteristics to novel scores. Therefore, the aim of this study was to compare the accuracy of qSOFA as an early warning score with SIRS, MEWS, and NEWS in patients with suspected infection on the wards and in the ED for predicting adverse outcomes.

Methods

Study Population

All adult patients admitted to the University of Chicago, an urban tertiary care medical center with approximately 500 beds, from November 2008 until January 2016 were eligible for inclusion in this observational study. Patients without vital sign or laboratory data documented in the ED or wards were excluded. In addition, patients who received mechanical ventilation or vasopressor medications before the first suspicion of infection were excluded because a decision support tool would not offer additional value for these patients because they would be admitted directly to the ICU. The protocol was approved by the University of Chicago Institutional Review Board (IRB #15-1705).

Data Collection

All time- and location-stamped vital signs, laboratory, orders (e.g., blood cultures and medications) and demographic data from the electronic health record for all admitted patients were collected by the University of Chicago's Clinical Research Data Warehouse, de-identified, and then made available on a secure SQL server for analysis. Non-physiologic values were changed to missing, as previously described (6).

Defining Infection

The definition from the original qSOFA paper by Seymour and colleagues was used to define the time of initial suspicion of infection (9). Specifically, suspected infection was defined as either (*1*) any culture order followed by an intravenous (IV) antibiotic within 72 hours or (*2*) an IV antimicrobial followed by a culture order within 24 hours. The time of the culture order or IV antimicrobial administration was denoted as the time of suspicion of infection, whichever came first. Medications were reviewed by three of the authors (S.S., N.P., and X.H.) to exclude prophylactic antibiotics. In addition, we excluded oral medications from the definition of suspicion of infection because IV antibiotics are recommended as the initial treatment of sepsis (14). Only patients who first met the suspicion of infection definition on the wards or ED were included in the study.

Outcomes

The primary outcome of the study was in-hospital mortality, and the secondary outcome was the composite of death or ICU stay at any point after a patient met the suspicion of infection criteria. A sensitivity analysis was performed by limiting the patient population to patients who met *International Classification of Diseases, 9th Revision, Clinical Modification* criteria for sepsis according to the criteria by Angus and colleagues (15). A second sensitivity analysis was performed by calculating accuracy for presumed septic shock, as defined by being both positive for the criteria by Angus and colleagues and receiving vasopressors after the time of suspicion of infection.

Sepsis-focused Criteria and General Early Warning Scores

The accuracy of two sepsis-focused criteria, SIRS and qSOFA, and two general early warning scores, MEWS and NEWS, were investigated in this study. SIRS criteria were defined as respiratory rate more than 20 breaths/min, temperature more than 38° C or less than 36° C, heart rate more than 90 beats/min, and white blood cell count more than 12,000/mm^3, less than 4,000/mm^3, or more than 10% bands (4). The qSOFA criteria were defined as systolic blood pressure ≤100 mm Hg, respiratory rate ≥22 breaths/min, and altered mental status (defined as either a Glasgow Coma Scale score ≤13 or an Alert Voice Pain Unresponsive scale other than "Alert") (9). MEWS and NEWS were calculated based

Table 1. Patient Characteristics

Characteristic	All Patients (N = 30,677)	Ward Patients (n = 12,154)	ED patients (n = 18,523)	P Value
Age, mean ± SD, yr	58 ± 18.0	57 ± 16.7	58 ± 18.9	<0.001
Female sex, n (%)	16,116 (53)	5,856 (48)	10,260 (55)	<0.001
Race, n (%)				<0.001
Black	17,813 (58)	4,384 (36)	13,429 (73)	
White	10,685 (35)	6,631 (55)	4,054 (22)	
Other	1,253 (4)	595 (5)	658 (4)	
Unknown	926 (3)	544 (4)	382 (2)	
LOS before time of suspicion, median (IQR), h	2.9 (1.1–7.9)	7.4 (2.4–30.5)	1.9 (0.8–4.3)	<0.001
LOS after time of suspicion, median (IQR), d	7.3 (5.8–11.6)	8.3 (6.0–14.3)	6.8 (5.8–10.1)	<0.001
Met Angus sepsis criteria, n (%)	8,744 (29)	3,350 (28)	5,394 (29)	0.003
Ever ICU transfer, n (%)	7,258 (24)	2,390 (20)	4,868 (26)	<0.001
Ever received vasopressor, n (%)	2,724 (9)	1,113 (9)	1,611 (9)	0.166
In-hospital mortality, n (%)	1,649 (5)	729 (6)	920 (5)	<0.001
Composite outcome, n (%)	7,385 (24)	2,385 (20)	5,000 (27)	<0.001

Definition of abbreviations: ED = emergency department; ICU = intensive care unit; IQR = interquartile range; LOS = length of stay.

on previously published tables (12, 13). Of note, our hospital has been collecting Glasgow Coma Scale data on ward patients since 2011, and before that period the components of Alert Voice Pain Unresponsive scale were documented.

Statistical Analysis

Patient characteristics were compared between those who first met the suspicion of infection definition on the wards versus ED using *t* tests, Wilcoxon rank sum tests, and χ^2 tests as necessary based on the distribution of the data. The highest value of each algorithm was calculated from the same contiguous non-ICU location segment (i.e., ED and ward) as when a patient first met the suspicion of the infection definition. Previous values were pulled forward if they were missing, and if no previous values were available, a median (normal) value was imputed, as per previous studies (9, 16, 17). Accuracy comparisons were performed using sensitivity, specificity, and the area under the receiver-operating characteristic curve was used to compare algorithm discrimination. A two-tailed $P < 0.05$ was considered statistically significant. Analyses were performed using Stata (version 14.1; StataCorp, College Station, TX).

Results

Study Population

A total of 445,073 patient records were available during the study period, of which 150,288 admissions occurred that had vital signs or laboratory values documented in the ED or wards and were eligible for inclusion in the study. Of these patients, 59,078 (39%) had at least one culture order, and 54,367 (36%) received at least one dose of IV antibiotics (*see* Figure E1 in the online supplement). The final study cohort consisted of 30,677 patients who met the definition of suspicion of infection outside the ICU, with both antibiotics and cultures within the predefined time window. Sixty percent (n = 18,523) first met this definition in the ED and 40% (n = 12,154) first met the definition on the wards.

Compared with patients with first suspicion of infection in the wards, ED patients were more likely to be female (55% vs. 48%; $P < 0.001$), black (73% vs. 36%; $P < 0.001$), meet the Angus sepsis criteria (29% vs. 28%; $P < 0.01$), and be admitted to the ICU at any point during their hospitalization (26% vs. 20%; $P < 0.01$). In addition, patients who first met suspicion criteria on the wards had higher in-hospital mortality (6% vs. 5%; $P < 0.01$) and length of stay after suspicion of infection (8.3 days vs. 6.8 days; $P < 0.01$) (Table 1). A total of 1,649
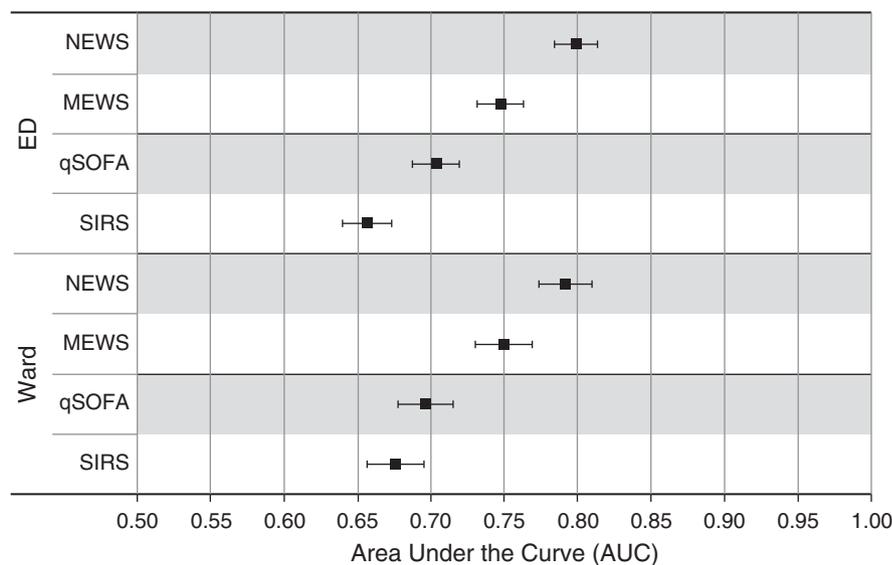
Figure 1. Discrimination of the different algorithms for predicting in-hospital mortality using each patient's highest score by location (*solid squares* represent point estimates, and *error bars* represent 95% confidence intervals). ED = emergency department; MEWS = Modified Early Warning Score; NEWS = National Early Warning Score; qSOFA = quick Sepsis-related Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

(5.4%) patients died, and 7,120 (23.2%) experienced an ICU stay after meeting suspicion of infection criteria. For those who experienced the composite outcome (death or ICU stay; n = 7,385), the median time to the outcome after time of first suspicion of infection was 14 hours (interquartile range, 6–66 h), and 71% of patients who experienced the composite outcome did so within 48 hours.

Score Distributions

Distributions of the highest SIRS criteria, qSOFA criteria, MEWS, and NEWS during the same contiguous non-ICU segment when a patient met suspicion of infection criteria are shown in Figures E2 to E5. Eighty-eight percent (n = 27,097) of patients met at least two SIRS criteria, and 38% (n = 11,729) met at least two qSOFA criteria during this time period. At the time of suspicion of infection, 51% of patients had met ≥2 SIRS, 9% had met ≥2, and 48% had met ≥1 qSOFA criteria at least once. The most common initial ≥2 qSOFA combination met by this time point was respiratory and blood pressure criteria in 5,016 (42.8%) patients, followed by respiratory rate and mental status criteria in 3,249 (27.7%) patients, blood pressure and mental status criteria in 1,993 (17.0%) patients, and all three criteria in 1,471 (12.5%) patients.

Accuracy Comparisons

Using each patient's highest score during their non-ICU stay, algorithm discrimination for in-hospital mortality in all non-ICU patients was highest for NEWS (area under the curve [AUC], 0.77; 95% confidence interval [CI], 0.76–0.79), followed by MEWS (AUC, 0.73; 95% CI, 0.71–0.74), the qSOFA score (AUC, 0.69; 95% CI, 0.67–0.70), and lowest for SIRS (AUC, 0.65; 95% CI, 0.63–0.66) ($P < 0.01$ for all pairwise comparisons). The relationship among the scores was consistent when comparing the ward and ED subgroups, with the AUCs being slightly lower on the wards (Figure 1). Furthermore, AUCs were similar but slightly lower for the composite outcome (see Figure E6). Using a patient's highest non-ICU score, NEWS ≥9 had a 72% sensitivity for in-hospital mortality compared with 71% for MEWS ≥5, 69% for qSOFA ≥2, and 94% for ≥2 SIRS (Table 2). Positive and negative predictive values at different thresholds for each

Table 2. Accuracy for the Outcomes across Different Score Thresholds Using the Highest Non–Intensive Care Unit Score for Each Patient

Score/Threshold	Mortality		Mortality or ICU Transfer	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
SIRS				
≥1	98.9	1.2	98.6	1.1
≥2*	93.8	12.3	91.0	13.0
≥3	77.5	43.8	67.7	45.9
≥4	36.8	84.0	26.1	85.7
qSOFA				
≥1	95.2	11.9	92.9	12.9
≥2*	68.7	63.5	53.6	66.7
≥3	19.0	96.0	10.4	97.0
MEWS				
≥1	100.0	0.0	100.0	0.0
≥2	96.2	5.6	96.2	6.0
≥3	92.8	20.3	89.2	22.3
≥4	84.7	44.1	76.1	48.4
≥5*	71.4	65.0	59.1	70.1
≥6	52.7	81.0	40.2	85.3
≥7	31.3	91.6	22.2	94.4
≥8	17.5	96.8	11.1	98.3
≥9	8.1	99.1	4.4	99.7
NEWS				
≥1	97.0	1.9	98.4	2.1
≥2	96.9	2.9	97.6	3.0
≥3	96.5	6.2	96.2	6.8
≥4	95.5	13.2	93.3	14.7
≥5†	95.1	15.0	92.6	16.7
≥5	93.6	23.1	89.2	25.8
≥6	91.0	34.6	84.0	38.7
≥7*	86.6	47.5	76.5	52.7
≥8	79.9	60.0	66.5	65.6
≥9	71.9	72.2	54.4	77.6
≥10	59.2	82.3	41.4	86.9
≥11	46.7	89.5	29.7	93.0
≥12	32.9	94.2	19.3	96.6
≥13	21.6	97.3	11.0	98.6
≥14	12.1	98.8	5.8	99.5
≥15	6.3	99.5	2.7	99.8

Definition of abbreviations: ICU = intensive care unit; MEWS = Modified Early Warning Score; NEWS = National Early Warning Score; qSOFA = quick Sepsis-related Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.
*Commonly used cutoff thresholds.
†NEWS total score ≥5 or at least one individual parameter score of 3.

score are found in Table E1 in the online supplement. Using the NEWS at a cutoff ≥9 would correctly reclassify 3% of patients who died and 9% of patients who did not die compared with using qSOFA ≥2. Furthermore, NEWS ≥8 would correctly reclassify 13% of patients who died or were transferred to the ICU compared with qSOFA ≥2 at a similar specificity. MEWS ≥5 would correctly reclassify 5% of patients who died or were transferred to the ICU and 3% of patients who did not experience the composite outcome compared with qSOFA ≥2. Figures E7 and E8 show the percentage of positive screens in the study population as

a function of sensitivity for each of the four tools for each outcome, which demonstrates the relative efficiency of NEWS and lack of efficiency for SIRS across the continuum.

Using a patient's highest score up until the time of first suspicion of infection, ≥9 NEWS had a sensitivity of 18% for the combined outcome in all non-ICU patients compared with 26% for ≥5 MEWS, 17% for ≥2 qSOFA, 65% for ≥1 qSOFA, and 62% for ≥2 SIRS. Most patients met SIRS criteria 17 hours before ICU transfer or death, compared with 12 hours for NEWS ≥7, and 5 hours for ≥2 and 17 hours for ≥1 qSOFA criteria (Figure 2).
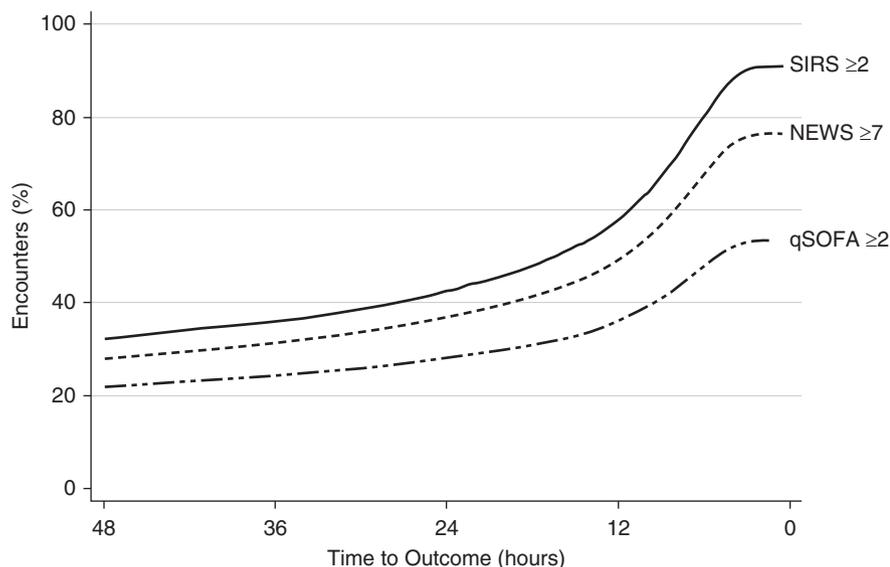
Figure 2. Cumulative percentage of patients meeting ≥2 quick Sepsis-related Organ Failure (qSOFA) criteria, ≥7 National Early Warning Score (NEWS) criteria, or ≥2 Systemic Inflammatory Response Syndrome (SIRS) criteria in the 48 hours before the composite outcome.

Sensitivity Analyses

The ranking of the scores was similar in the cohort of patients who met the Angus *International Classification of Diseases, 9th edition* criteria for sepsis (n = 8,744), with AUCs for in-hospital mortality of 0.71 (95% CI, 0.69–0.72) for NEWS, 0.66 (95% CI, 0.64–0.68) for MEWS, 0.63 (95% CI, 0.61–0.64) for qSOFA, and 0.60 (95% CI, 0.58–0.62) for SIRS ($P < 0.01$ for all pairwise comparisons). In this population, NEWS ≥7 had a sensitivity of 82% for the composite outcome compared with a sensitivity of 63% for qSOFA ≥2 and 93% for ≥2 SIRS criteria. Accuracy was higher for the subset of patients who also received vasopressor drugs after the onset of infection, with a sensitivity of 87% for NEWS ≥7, 70% for ≥2 qSOFA, and 93% for ≥2 SIRS criteria for the composite outcome.

Discussion

In this observational cohort study, we found that although qSOFA was more accurate than SIRS for predicting in-hospital mortality and ICU transfer in both ward and ED patients, it was less accurate than the general early warning scores. NEWS was the most accurate tool for predicting adverse outcomes in both ED and ward patients. This is important because early warning scores such as the

MEWS are in widespread use in Europe and the United States, and the NEWS is now mandated in the United Kingdom as a tool to identify patients outside the ICU at high risk of clinical deterioration (13, 18–20).

The SIRS criteria have been part of the sepsis definition for more than two decades, and have been criticized in the literature for almost as long (3, 5). In particular, they have been shown to have poor specificity, with up to 90% of ICU patients and almost one-half of ward patients meeting at least two of four criteria at some point in their stay (6, 7). These findings and others led to the recent update of the sepsis definitions in 2016 by the Society of Critical Care Medicine/European Society of Intensive Care Medicine task force (8). The qSOFA score was published with these updated sepsis definitions as a tool that could be used outside the ICU to prompt clinicians to consider possible sepsis (9). Although the new sepsis definitions were endorsed by several societies, other societies, including Chest and the American College of Emergency Physicians, did not endorse them (11). One criticism of these new criteria has been that they will identify patients too late in their course, after organ dysfunction has already occurred (11). Furthermore, some authors have questioned the use of qSOFA as a clinical decision tool because of how it was developed (21).

We found that less than one in five patients who later go on to die or be transferred to the ICU will have met ≥2 qSOFA criteria by the time of infection suspicion, which illustrates the importance of score recalculation after the initiation of therapy. Furthermore, most patients who experienced the composite outcome met ≥2 SIRS criteria more than 17 hours before the composite outcome compared with only 5 hours for ≥2 qSOFA, with almost one-half of patients still not meeting ≥2 qSOFA criteria at the time of the outcome. The previous paradigm using SIRS favored higher sensitivity at the cost of specificity, whereas using a qSOFA score cutoff of ≥2 increased specificity at the cost of sensitivity. Interestingly, we found that using ≥1 qSOFA criteria had similar accuracy and timing compared with ≥2 SIRS criteria, which suggests that this lower threshold could be used if an earlier and more sensitive cutoff were desired.

Our study found that general early warning scores are more accurate than qSOFA for predicting adverse outcomes in the ED and on the wards. The MEWS, and its derivatives, such as the NEWS, which was endorsed by the Royal College of Physicians for standard use across the United Kingdom, are already used in many hospitals for Rapid Response System activation (13, 18, 19). We found NEWS to be the most accurate score we studied. qSOFA has the advantage of simplicity, and errors in manually calculating general early warning scores have been described (22). However, the NEWS and MEWS offer several more thresholds to vary sensitivity and specificity to resource availability. In addition, these early warning scores are increasingly becoming available for automated calculation within the electronic health record. Furthermore, data suggest that general early warning scores such as MEWS and the electronic Cardiac Arrest Risk Triage model add useful predictive information to clinical judgment (23–25). Currently, it is unknown whether using ≥2 qSOFA criteria, which would identify a patient with both mental status changes and hypotension as a patient to pay more attention to, offers additional value above caregiver intuition. Overall, our study provides evidence that hospitals already using the NEWS or MEWS would not benefit from switching to qSOFA for use as an early warning score because of the costs

and risks of retraining caregivers to use a new scoring system.

Our study had several limitations. First, this was a single-center investigation in an academic U.S. hospital; therefore, the results may not be generalizable to other settings. In addition, there is no gold standard to determine when a patient is infected, so we might have excluded patients who were infected and included others who were not. However, we used the same definition as the original qSOFA study and also included only IV medications, which should improve the

validity of our results. Finally, we only studied 2 of the more than 100 published early warning scores in the literature (26). Because most of these scores are similar to the MEWS and NEWS, and these are two of the most highly cited scores in common use, we believed that these results would be of value for clinicians.

In conclusion, we found that general early warning scores were more accurate than the qSOFA score for predicting in-hospital mortality and ICU transfer in non-ICU patients with suspicion of infection, with the NEWS being the most

accurate score in our study. These findings have important implications for clinicians at the bedside, hospitals, and countries implementing these scoring systems in practice. ■

References

1. Liu V, Escobar GJ, Greene JD, Soule J, Whippy A, Angus DC, Iwashyna TJ. Hospital deaths in patients with sepsis from 2 independent cohorts. *JAMA* 2014;312:90–92.
2. Lagu T, Rothberg MB, Shieh MS, Pekow PS, Steingrub JS, Lindenauer PK. Hospitalizations, costs, and outcomes of severe sepsis in the United States 2003 to 2007. *Crit Care Med* 2012;40:754–761.
3. Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, Schein RM, Sibbald WJ. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest* 1992;101:1644–1655.
4. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, Cohen J, Opal SM, Vincent JL, Ramsay G; SCCM/ESICM/ACCP/ATS/SIS. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. *Crit Care Med* 2003;31:1250–1256.
5. Vincent JL. Dear SIRS, I'm sorry to say that I don't like you.... *Crit Care Med* 1997;25:372–374.
6. Churpek MM, Zadravecz FJ, Winslow C, Howell MD, Edelson DP. Incidence and prognostic value of the systemic inflammatory response syndrome and organ dysfunctions in ward patients. *Am J Respir Crit Care Med* 2015;192:958–964.
7. Sprung CL, Sakr Y, Vincent JL, Le Gall JR, Reinhart K, Ranieri VM, Gerlach H, Fielden J, Groba CB, Payen D. An evaluation of systemic inflammatory response syndrome signs in the Sepsis Occurrence In Acutely Ill Patients (SOAP) study. *Intensive Care Med* 2006;32:421–427.
8. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, *et al*. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:801–810.
9. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, Rubenfeld G, Kahn JM, Shankar-Hari M, Singer M, *et al*. Assessment of clinical criteria for sepsis: for the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:762–774.
10. Levy MM, Rhodes A, Phillips GS, Townsend SR, Schorr CA, Beale R, Osborn T, Lemeshow S, Chiche JD, Artigas A, *et al*. Surviving sepsis campaign: association between performance metrics and outcomes in a 7.5-year study. *Crit Care Med* 2015;43:3–12.
11. Simpson SQ. New sepsis criteria: a change we should not make. *Chest* 2016;149:1117–1118.
12. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified early warning score in medical admissions. *QJM* 2001;94:521–526.
13. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013;84:465–470.
14. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, Sevransky JE, Sprung CL, Douglas IS, Jaeschke R, *et al*.; Surviving Sepsis Campaign Guidelines Committee including the Pediatric Subgroup. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. *Crit Care Med* 2013;41:580–637.
15. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Crit Care Med* 2001;29:1303–1310.
16. Churpek MM, Yuen TC, Winslow C, Robicsek AA, Meltzer DO, Gibbons RD, Edelson DP. Multicenter development and validation of a risk stratification tool for ward patients. *Am J Respir Crit Care Med* 2014; 190:649–655.
17. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A, *et al*. The APACHE III prognostic system: risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100: 1619–1636.
18. Royal College of Physicians. National early warning score (NEWS): standardising the assessment of acute illness severity in the NHS. London: Royal College of Physicians; 2012.
19. Jones DA, DeVita MA, Bellomo R. Rapid-response teams. *N Engl J Med* 2011;365:139–146.
20. Smith ME, Chiovaro JC, O'Neil M, Kansagara D, Quiñones AR, Freeman M, Motu'apuaka ML, Slatore CG. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. *Ann Am Thorac Soc* 2014;11:1454–1465.
21. Moskowitz A, Andersen LW, Cocchi M, Donnino MW. The misapplication of severity-of-illness scores toward clinical decision making. *Am J Respir Crit Care Med* 2016;194:256–258.
22. Prytherch DR, Smith GB, Schmidt P, Featherstone PI, Stewart K, Knight D, Higgins B. Calculating early warning scores: a classroom comparison of pen and paper and hand-held computer methods. *Resuscitation* 2006;70:173–178.
23. Patel AR, Zadravecz FJ, Young RS, Williams MV, Churpek MM, Edelson DP. The value of clinical judgment in the detection of clinical deterioration. *JAMA Intern Med* 2015;175:456–458.
24. Fullerton JN, Price CL, Silvey NE, Brace SJ, Perkins GD. Is the Modified Early Warning Score (MEWS) superior to clinician judgement in detecting critical illness in the pre-hospital environment? *Resuscitation* 2012;83:557–562.
25. Kang MA, Churpek MM, Zadravecz FJ, Adhikari R, Twu NM, Edelson DP. Real-time risk prediction on the wards: a feasibility study. *Crit Care Med* 2016;44:1468–1473.
26. Churpek MM, Yuen TC, Edelson DP. Risk stratification of hospitalized patients on the wards. *Chest* 2013;143:1758–1765.