

Identifying the Sickest During Triage: Using Point-of-Care Severity Scores to Predict Prognosis in Emergency Department Patients With Suspected Sepsis

Priya A Prasad, PhD, MPH^{1*}, Margaret C Fang, MD, MPH¹, Sandra P Martinez, MPH¹, Kathleen D Liu, MD, PhD², Kirsten N Kangelaris, MD, MAS¹

¹Division of Hospital Medicine, Department of Medicine, University of California, San Francisco, California; ²Divisions of Nephrology and Critical Care Medicine, Department of Medicine, University of California, San Francisco, San Francisco, California.

BACKGROUND: Sepsis progresses rapidly and is associated with considerable morbidity and mortality. Bedside risk stratification scores can quickly identify patients at greatest risk of poor outcomes; however, there is lack of consensus on the best scale to use.

OBJECTIVE: To compare the ability of quick Sequential Organ Failure Assessment (qSOFA), the National Early Warning System (NEWS2), and the Shock Index—which does not require mental status assessment—to predict poor outcomes among patients with suspected sepsis during triage.

DESIGN, SETTING, AND PARTICIPANTS: Retrospective cohort study of adults presenting to an academic emergency department (ED) from June 2012 to December 2018 who had blood cultures and intravenous antibiotics within 24 hours.

MAIN OUTCOMES AND MEASURES: Clinical data were collected from the electronic health record. Patients were considered positive at qSOFA ≥ 2 , Shock Index > 0.7 , or NEWS2 ≥ 5 scores. We calculated test characteristics and area under the receiver operating characteristics curves

(AUROCs) to predict in-hospital mortality and ED-to-intensive care unit (ICU) admission.

RESULTS: We included 23,837 ED patients; 1,921 (8.1%) were qSOFA-positive, 4,273 (17.9%) Shock Index-positive, and 11,832 (49.6%) NEWS2-positive. There were 1,427 (6.0%) deaths and 3,149 (13.2%) ED-to-ICU admissions in the sample. NEWS2 had the highest sensitivity for in-hospital mortality (76.0%) and ED-to-ICU admission (78.9%). qSOFA had the highest specificity for in-hospital mortality (93.4%) and ED-to-ICU admission (95.2%). Shock Index exhibited the highest AUROC for in-hospital mortality (0.648; 95% CI, 0.635-0.662) and ED-to-ICU admission (0.680; 95% CI, 0.617-0.689). Test characteristics were similar among those with sepsis.

CONCLUSIONS: Institution priorities should drive score selection, balancing sensitivity and specificity. In our study, qSOFA was highly specific and NEWS2 was the most sensitive for ruling out patients at high risk. Performance of the Shock Index fell between qSOFA and NEWS2 and could be considered because it is easy to implement. *Journal of Hospital Medicine* 2021;16:453-461. © 2021 Society of Hospital Medicine

Sepsis is the leading cause of in-hospital mortality in the United States.¹ Sepsis is present on admission in 85% of cases, and each hour delay in antibiotic treatment is associated with 4% to 7% increased odds of mortality.^{2,3} Prompt identification and treatment of sepsis is essential for reducing morbidity and mortality, but identifying sepsis during triage is challenging.²

Risk stratification scores that rely solely on data readily available at the bedside have been developed to quickly identify those at greatest risk of poor outcomes from sepsis in real time. The quick Sequential Organ Failure Assessment (qSOFA) score, the National Early Warning System (NEWS2), and the Shock Index are easy-to-calculate measures that use

routinely collected clinical data that are not subject to laboratory delay. These scores can be incorporated into electronic health record (EHR)-based alerts and can be calculated longitudinally to track the risk of poor outcomes over time. qSOFA was developed to quantify patient risk at bedside in non-intensive care unit (ICU) settings, but there is no consensus about its ability to predict adverse outcomes such as mortality and ICU admission.⁴⁻⁶ The United Kingdom's National Health Service uses NEWS2 to identify patients at risk for sepsis.⁷ NEWS has been shown to have similar or better sensitivity in identifying poorer outcomes in sepsis patients compared with systemic inflammatory response syndrome (SIRS) criteria and qSOFA.^{4,8-11} However, since the latest update of NEWS2 in 2017, there has been little study of its predictive ability. The Shock Index is a simple bedside score (heart rate divided by systolic blood pressure) that was developed to detect changes in cardiovascular performance before systemic shock onset. Although it was not developed for infection and has not been regularly applied in the sepsis literature, the Shock Index might be useful for identifying patients at increased risk

*Corresponding Author: Priya A Prasad, PhD, MPH; Priya.Prasad@ucsf.edu.

Published online first July 21, 2021.

Find additional supporting information in the online version of this article.

Received: October 7, 2020; Revised: April 20, 2021; Accepted: April 27, 2021

© 2021 Society of Hospital Medicine DOI 10.12788/jhm.3642

TABLE 1. Data Elements Collected in the 72 Hours After Emergency Department Presentation to Identify Patients Meeting qSOFA, Shock Index, and NEWS2 Criteria

qSOFA (range, 0-3)	Shock Index (heart rate/systolic blood pressure)	NEWS2 (range, 0-20)
Systolic blood pressure ≤ 100 mmHg	Systolic blood pressure	Systolic blood pressure
	Heart rate	Heart rate
Respiratory rate ≥ 22		Respiratory rate
Glasgow Coma Scale < 15		Glasgow Coma Scale value ^a
		Oxygen saturation
		Requirement for supplemental oxygen
		Temperature

Abbreviations: NEWS2, National Early Warning System; qSOFA, quick Sequential Organ Failure Assessment.

^a NEWS2 uses level of consciousness, which is not collected in our electronic health record. We used the Glasgow Coma Scale value as a surrogate for level of consciousness.

of poor outcomes. Patients with higher and sustained Shock Index scores are more likely to experience morbidity, such as hyperlactatemia, vasopressor use, and organ failure, and also have an increased risk of mortality.¹²⁻¹⁴

Although the predictive abilities of these bedside risk stratification scores have been assessed individually using standard binary cut-points, the comparative performance of qSOFA, the Shock Index, and NEWS2 has not been evaluated in patients presenting to an emergency department (ED) with suspected sepsis. Our objective was to provide a head-to-head comparison of the test characteristics of qSOFA, the Shock Index, and NEWS2 calculated at ED triage for predicting in-hospital mortality and ED-to-ICU admission in patients with suspected sepsis to help health systems and providers select screening measures.

METHODS

Design and Setting

We conducted a retrospective cohort study of ED patients who presented with suspected sepsis to the University of California San Francisco (UCSF) Helen Diller Medical Center at Parnassus Heights between June 1, 2012, and December 31, 2018. Our institution is a 785-bed academic teaching hospital with approximately 30,000 ED encounters per year. The study was approved with a waiver of informed consent by the UCSF Human Research Protection Program.

Participants

We use an Epic-based EHR platform (Epic 2017, Epic Systems Corporation) for clinical care, which was implemented on June 1, 2012. All data elements were obtained from Clarity, the relational database that stores Epic's inpatient data. The study included encounters for patients age ≥ 18 years who had blood cultures ordered within 24 hours of ED presentation and administration of intravenous antibiotics within 24 hours. Repeat encounters were treated independently in our analysis.

Outcomes and Measures

We compared the ability of qSOFA, the Shock Index, and NEWS2 to predict in-hospital mortality and admission to the ICU from the ED (ED-to-ICU admission). We used the most abnormal vital signs and clinical assessments gathered within 30 minutes of ED presentation to identify patients who were qSOFA-positive, Shock Index-positive, and NEWS2-positive based on standard cut-points of risk. Data elements used to calculate qSOFA, Shock Index, and NEWS2 included blood pressure, heart rate, respiratory rate, Glasgow Coma Scale (GCS) score, oxygen saturation, requirement for supplemental oxygen, and temperature (Table 1). Patients were considered positive if they had a qSOFA score ≥ 2 , Shock Index of > 0.7 , and NEWS2 ≥ 5 based on triage vital signs.^{7,15,16} We considered patients to have altered mental status, a criterion used for NEWS2, if they had a GCS score < 15 instead of using the "alert, verbal, confusion, pain, unresponsive" scale, which is not captured in our EHR, a method that has been used in earlier studies.^{17,18} Missing assessments were considered normal. Although our primary analysis focused on the scores calculated within 30 minutes of ED presentation, we performed a sensitivity analysis examining scores calculated within 1 hour of ED presentation in the event of a delay in gathering triage vital sign data.

We compared demographic and clinical characteristics of patients who were positive for qSOFA, the Shock Index, and NEWS2. Demographic data were extracted from the EHR and included primary language, age, sex, and insurance status. All *International Classification of Diseases* (ICD)-9/10 diagnosis codes were pulled from Clarity billing tables. We used the Elixhauser comorbidity groupings¹⁹ of ICD-9/10 codes present on admission to identify preexisting comorbidities and underlying organ dysfunction. To estimate burden of comorbid illnesses, we calculated the validated van Walraven comorbidity index,²⁰ which provides an estimated risk of in-hospital death based on documented Elixhauser comorbidities. Admission level of care (acute, stepdown, or intensive care) was collected for inpatient admissions to assess initial illness severity.²¹ We also evaluated

discharge disposition and in-hospital mortality. Index blood culture results were collected, and dates and timestamps of mechanical ventilation, fluid, vasopressor, and antibiotic administration were obtained for the duration of the encounter.

UCSF uses an automated, real-time, algorithm-based severe sepsis alert that is triggered when a patient meets ≥ 2 SIRS criteria and again when the patient meets severe sepsis or septic shock criteria (ie, ≥ 2 SIRS criteria in addition to end-organ dysfunction and/or fluid nonresponsive hypotension). This sepsis screening alert was in use for the duration of our study.²²

Statistical Analysis

We performed a subgroup analysis among those who were diagnosed with sepsis, according to the 2016 Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) criteria. Sepsis is defined as a change in Sequential Organ Failure Assessment (SOFA) score of ≥ 2 points within the first 48 hours.²³ Additionally, patients meeting Sepsis-3 criteria needed to (1) receive ≥ 4 days of sequential antibiotic therapy or experience death or discharge to hospice before 4 days of antibiotic therapy or (2) have a validated sepsis discharge billing code. These parameters were added to increase the specificity of our sample.²⁴

All statistical analyses were conducted using Stata 14 (StataCorp). We summarized differences in demographic and clinical characteristics among the populations meeting each severity score but elected not to conduct hypothesis testing because patients could be positive for one or more scores. We calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each score to predict in-hospital mortality and ED-to-ICU admission. To allow comparison with other studies, we also created a composite outcome of either in-hospital mortality or ED-to-ICU admission. To assess score discrimination to predict in-hospital mortality and ED-to-ICU admission, we calculated the area under the receiver operating characteristic curve (AUROC) along with asymptotic normal 95% CI using the "roctab" command considering a binary cut-point, as well as the full range of scores measured in the cohort. The AUROC range from 0.50 to 1.00 and a score in the 0.70 to 0.80 range can be considered fair.²⁵ We assessed significant differences between severity score AUROCs using the DeLong method²⁶ implemented through Stata 14's "roccomp" command. As a sensitivity analysis, we explored whether the standard cut-points for qSOFA, the Shock Index, and NEWS2 provided the highest AUROC in our population by calculating test characteristics for several score cut-points.

RESULTS

Within our sample 23,837 ED patients had blood cultures ordered within 24 hours of ED presentation and were considered to have suspected sepsis. The mean age of the cohort was 60.8 years, and 1,612 (6.8%) had positive blood cultures. A total of 12,928 patients (54.2%) were found to have sepsis. We documented 1,427 in-hospital deaths (6.0%) and 3,149 (13.2%) ED-to-ICU admissions. At ED triage 1,921 (8.1%) were

qSOFA-positive, 4,273 (17.9%) were Shock Index-positive, and 11,832 (49.6%) were NEWS2-positive. At ED triage, blood pressure, heart rate, respiratory rate, and oxygen saturated were documented in $>99\%$ of patients, 93.5% had temperature documented, and 28.5% had GCS recorded. If the window of assessment was widened to 1 hour, GCS was only documented among 44.2% of those with suspected sepsis.

Demographic Characteristics and Clinical Course

We identified significant differences when comparing demographic and clinical characteristics among patients who scored positive for the three severity measures at triage (Table 2). Although no hypothesis testing was conducted because patients could meet one or more scores, qSOFA-positive patients were older (median 70, 66, and 64 years, respectively), more likely to have Medicare as the primary payor (67.6% vs 59.7% vs 56.6%), to have chronic renal failure (26.1%, 23.1%, and 23.3%, respectively), to have a greater degree of underlying comorbidities based on the van Walraven Comorbidity Index (median 15, 12, and 11, respectively), and to be admitted to the ICU from the ED (48.1%, 36.3%, and 21.0%, respectively) compared with those positive for the Shock Index or NEWS2.

qSOFA-positive patients received antibiotics more quickly than those who were Shock Index-positive or NEWS2-positive (median 1.5, 1.8, and 2.8 hours after admission, respectively). In addition, those who were qSOFA-positive were more likely to have a positive blood culture (10.9%, 9.4%, and 8.5%, respectively) and to receive an EHR-based diagnosis of sepsis (77.0%, 69.6%, and 60.9%, respectively) than those who were Shock Index- or NEWS2-positive. Those who were qSOFA-positive also were more likely to be mechanically ventilated during their hospital stay (25.4%, 19.2%, and 10.8%, respectively) and to receive vasopressors (33.5%, 22.5%, and 12.2%, respectively). In-hospital mortality also was more common among those who were qSOFA-positive at triage (23.4%, 15.3%, and 9.2%, respectively).

Because both qSOFA and NEWS2 incorporate GCS, we explored baseline characteristics of patients with GCS documented at triage ($n = 6,794$). These patients were older (median age 63 and 61 years, $P < .0001$), more likely to be male (54.9% and 53.4%, $P = .0031$), more likely to have renal failure (22.8% and 20.1%, $P < .0001$), more likely to have liver disease (14.2% and 12.8%, $P = .006$), had a higher van Walraven comorbidity score on presentation (median 10 and 8, $P < .0001$), and were more likely to go directly to the ICU from the ED (20.2% and 10.6%, $P < .0001$). However, among the 6,397 GCS scores documented at triage, only 1,579 (24.7%) were abnormal.

Test Characteristics of qSOFA, Shock Index, and NEWS2 for Predicting In-hospital Mortality and ED-to-ICU Admission

Among 23,837 patients with suspected sepsis, NEWS2 had the highest sensitivity for predicting in-hospital mortality (76.0%; 95% CI, 73.7%-78.2%) and ED-to-ICU admission (78.9%; 95% CI, 77.5%-80.4%) but had the lowest specificity for in-hospital mortality (52.0%; 95% CI, 51.4%-52.7%) and for ED-to-ICU

TABLE 2. Demographic Characteristics and Clinical Course of Patients With Suspected Infection and Populations Who Met Standard Cut-Points of qSOFA, Shock Index, and NEWS2

Variable	Full cohort (N = 23,837)	qSOFA (n = 1,921)	Shock Index (n = 4,237)	NEWS2 (n = 11,832)
Characteristics at presentation				
Age, median (IQR), y	62 (48-75)	70 (57-83)	66 (53-79)	64 (50-78)
Male	12,817 (53.8)	1,084 (56.4)	2,404 (56.3)	6,377 (53.9)
Payor type				
Medicare	12,400 (52.0)	1,299 (67.6)	2,550 (59.7)	6,701 (56.6)
Medi-Cal	5,964 (25.0)	342 (17.8)	943 (22.1)	2,804 (23.7)
Commercial	5,038 (21.1)	261 (13.6)	730 (17.1)	2,195 (18.6)
Other/self-Pay	435 (1.8)	19 (1.0)	50 (1.2)	132 (1.1)
Chronic renal failure	4,970 (20.9)	501 (26.1)	988 (23.1)	2,751 (23.3)
Chronic liver disease	3,147 (13.2)	312 (16.2)	637 (14.9)	1,718 (14.5)
Cancer ^a				
Lymphoma	779 (3.3)	62 (3.2)	156 (3.7)	404 (3.4)
Solid tumor	2,526 (10.6)	223 (11.6)	517 (12.1)	1,390 (11.8)
Metastatic cancer	2,280 (9.6)	209 (10.9)	494 (11.6)	1,275 (10.8)
van Walraven Comorbidity Index value, median (IQR)	8 (1-16)	15 (7-22)	12 (5-20)	11 (4-19)
Admission level of care				
Discharged from ED	1,501 (6.3)	14 (0.5)	56 (1.3)	381 (3.2)
Admission to outside facility	446 (1.9)	30 (1.7)	65 (1.5)	202 (1.7)
Acute care	13,644 (57.2)	391 (20.4)	1,223 (28.6)	5,635 (47.6)
Stepdown care	5,097 (21.4)	563 (29.3)	1,379 (32.3)	3,128 (26.4)
Intensive care	3,149 (13.2)	923 (48.1)	1,550 (36.3)	2,486 (21.0)
Clinical course				
Time to antibiotics, median (IQR), h	3.2 (1.8-5.4)	1.5 (0.8-3.0)	1.8 (1.0-3.5)	2.8 (1.5-5.0)
Receipt of mechanical ventilation	1,719 (7.4)	488 (25.4)	820 (19.2)	1,283 (10.8)
Receipt of vasopressors	1,759 (7.4)	644 (33.5)	963 (22.5)	1,441 (12.2)
Positive blood culture	1,612 (6.8)	209 (10.9)	401 (9.4)	1,007 (8.5)
Received an EHR-based diagnosis of sepsis	12,928 (54.2)	1,480 (77.0)	2,974 (69.6)	7,208 (60.9)
In-hospital mortality	1,427 (6.0)	449 (23.4)	654 (15.3)	1,084 (9.2)

All values represent No. (%) unless otherwise noted.

Abbreviations: ED, emergency department; EHR, electronic health record; IQR, interquartile range; NEWS2, National Early Warning System; qSOFA, quick Sequential Organ Failure Assessment.

admission (54.8%; 95% CI, 54.1%-55.5%) (Table 3). qSOFA had the lowest sensitivity for in-hospital mortality (31.5%; 95% CI, 29.1%-33.9%) and ED-to-ICU admission (29.3%; 95% CI, 27.7%-30.9%) but the highest specificity for in-hospital mortality (93.4%; 95% CI, 93.1%-93.8%) and ED-to-ICU admission (95.2%; 95% CI, 94.9%-95.5%). The Shock Index had a sensitivity that fell between qSOFA and NEWS2 for in-hospital mortality (45.8%; 95% CI, 43.2%-48.5%) and ED-to-ICU admission (49.2%; 95% CI, 47.5%-51.0%). The specificity of the Shock Index also was between qSOFA and NEWS2 for in-hospital mortality (83.9%; 95% CI, 83.4%-84.3%) and ED-to-ICU admission (86.8%; 95% CI, 86.4%-87.3%). All three scores exhibited relatively low PPV, ranging from 9.2% to 23.4% for in-hospital mortality and 21.0%

to 48.0% for ED-to-ICU triage. Conversely, all three scores exhibited relatively high NPV, ranging from 95.5% to 97.1% for in-hospital mortality and 89.8% to 94.5% for ED-to-ICU triage. The patterns in sensitivity and specificity for in-hospital mortality and ED-to-ICU admission were similar among the 12,928 patients who received an EHR-based sepsis diagnosis with the tests generally demonstrating lower specificities, higher PPVs, and lower NPVs (Table 3).

When considering a binary cutoff, the Shock Index exhibited the highest AUROC for in-hospital mortality (0.648; 95% CI, 0.635-0.662) and had a significantly higher AUROC than qSOFA (AUROC, 0.625; 95% CI, 0.612-0.637; $P = .0005$), but there was no difference compared with NEWS2 (AUROC,

TABLE 3. Test Characteristics of Point-of-Care Severity Scores for Predicting Prognosis in ED Patients Presenting With Suspected Sepsis and Those With EHR-Based Sepsis Diagnosis

Suspected sepsis (n = 23,837)				
	Sensitivity (95% CI), %	Specificity (95% CI), %	PPV (95% CI), %	NPV (95% CI), %
In-hospital mortality				
qSOFA	31.5 (29.1-33.9)	93.4 (93.1-93.8)	23.4 (21.5-25.3)	95.5 (95.3-95.8)
Shock Index	45.8 (43.2-48.5)	83.9 (83.4-84.3)	15.3 (14.2-16.4)	96.0 (95.8-96.3)
NEWS2	76.0 (73.7-78.2)	52.0 (51.4-52.7)	9.2 (8.7-9.7)	97.1 (96.8-97.4)
ED-to-ICU Admission				
qSOFA	29.3 (27.7-30.9)	95.2 (94.9-95.5)	48.0 (45.4-50.3)	89.8 (89.4-90.2)
Shock Index	49.2 (47.5-51.0)	86.8 (86.4-87.3)	36.3 (34.8-37.7)	91.8 (91.4-92.2)
NEWS2	78.9 (77.5-80.4)	54.8 (54.1-55.5)	21.0 (20.3-21.8)	94.5 (94.1-94.9)
EHR-based sepsis diagnosis (n = 12,928)				
	Sensitivity (95% CI), %	Specificity (95% CI), %	PPV (95% CI), %	NPV (95% CI), %
In-hospital mortality				
qSOFA	32.7 (30.2-35.3)	91.0 (90.4-91.5)	29.1 (26.8-31.5)	92.3 (91.8-92.7)
Shock Index	47.6 (44.9-50.3)	79.8 (79.0-80.5)	21.1 (19.6-22.6)	93.1 (92.6-93.6)
NEWS2	76.2 (73.8-78.4)	46.6 (45.6-47.5)	13.9 (13.1-14.7)	94.5 (93.9-95.1)
ED-to-ICU Admission				
qSOFA	32.4 (30.5-34.3)	93.3 (92.8-93.8)	52.2 (49.6-54.8)	85.9 (85.2-86.5)
Shock Index	51.9 (49.8-53.9)	83.5 (82.8-84.2)	41.7 (39.9-43.5)	88.4 (87.8-89.1)
NEWS2	80.1 (78.4-81.7)	49.8 (48.8-50.7)	26.5 (25.5-27.6)	91.7 (90.9-92.4)

Abbreviations: ED, emergency department; EHR, electronic health record; ICU, intensive care unit; NEWS2, National Early Warning System; qSOFA, quick Sequential Organ Failure Assessment.

0.640; 95% CI, 0.628-0.652; $P = .2112$). NEWS2 had a significantly higher AUROC than qSOFA for predicting in-hospital mortality ($P = .0227$). The Shock Index also exhibited the highest AUROC for ED-to-ICU admission (0.680; 95% CI, 0.617-0.689), which was significantly higher than the AUROC for qSOFA ($P < .0001$) and NEWS2 ($P = 0.0151$). NEWS2 had a significantly higher AUROC than qSOFA for predicting ED-to-ICU admission ($P < .0001$). Similar findings were seen in patients found to have sepsis. When considering the range of possible scores measured in our cohort, qSOFA and NEWS2 exhibited higher AUROCs for in-hospital mortality and ED-to-ICU admission than the Shock Index among patients with suspected infection and the subgroup with a sepsis diagnosis (Figure). The AUROCs of the scores were statistically significantly different for both in-hospital mortality ($P = .0026$) and ED-to-ICU admission ($P < .0001$). As a sensitivity analysis, we varied the binary cut-points of qSOFA, the Shock Index, and NEWS2 and calculated test characteristics within our study cohort (Appendix). For our institution, the qSOFA cut-point with the highest AUROC would be qSOFA > 0 for both in-hospital mortality (AUROC, 0.699; 95% CI, 0.687-0.711) and ED-to-ICU admission (AUROC,

0.716; 95% CI, 0.707-0.724), with 36.5% of the cohort meeting qSOFA. The NEWS2 cut-point with the highest AUROC would be NEWS2 ≥ 7 for both in-hospital mortality (AUROC, 0.653; 95% CI, 0.640-0.666) and ED-to-ICU admission (AUROC, 0.677; 95% CI, 0.668-0.686), with 20.3% of the cohort meeting NEWS2 at this cut-point. The standard Shock Index cut-point ≥ 0.7 exhibited the highest AUROC for in-hospital mortality and ED-to-ICU admission at our institution.

DISCUSSION

In this retrospective cohort study of 23,837 patients who presented to the ED with suspected sepsis, the standard qSOFA threshold was met least frequently, followed by the Shock Index and NEWS2. NEWS2 had the highest sensitivity but the lowest specificity for predicting in-hospital mortality and ED-to-ICU admission, making it a challenging bedside risk stratification scale for identifying patients at risk of poor clinical outcomes. When comparing predictive performance among the three scales, qSOFA had the highest specificity and the Shock Index had the highest AUROC for in-hospital mortality and ED-to-ICU admission in this cohort of patients with suspected sepsis.

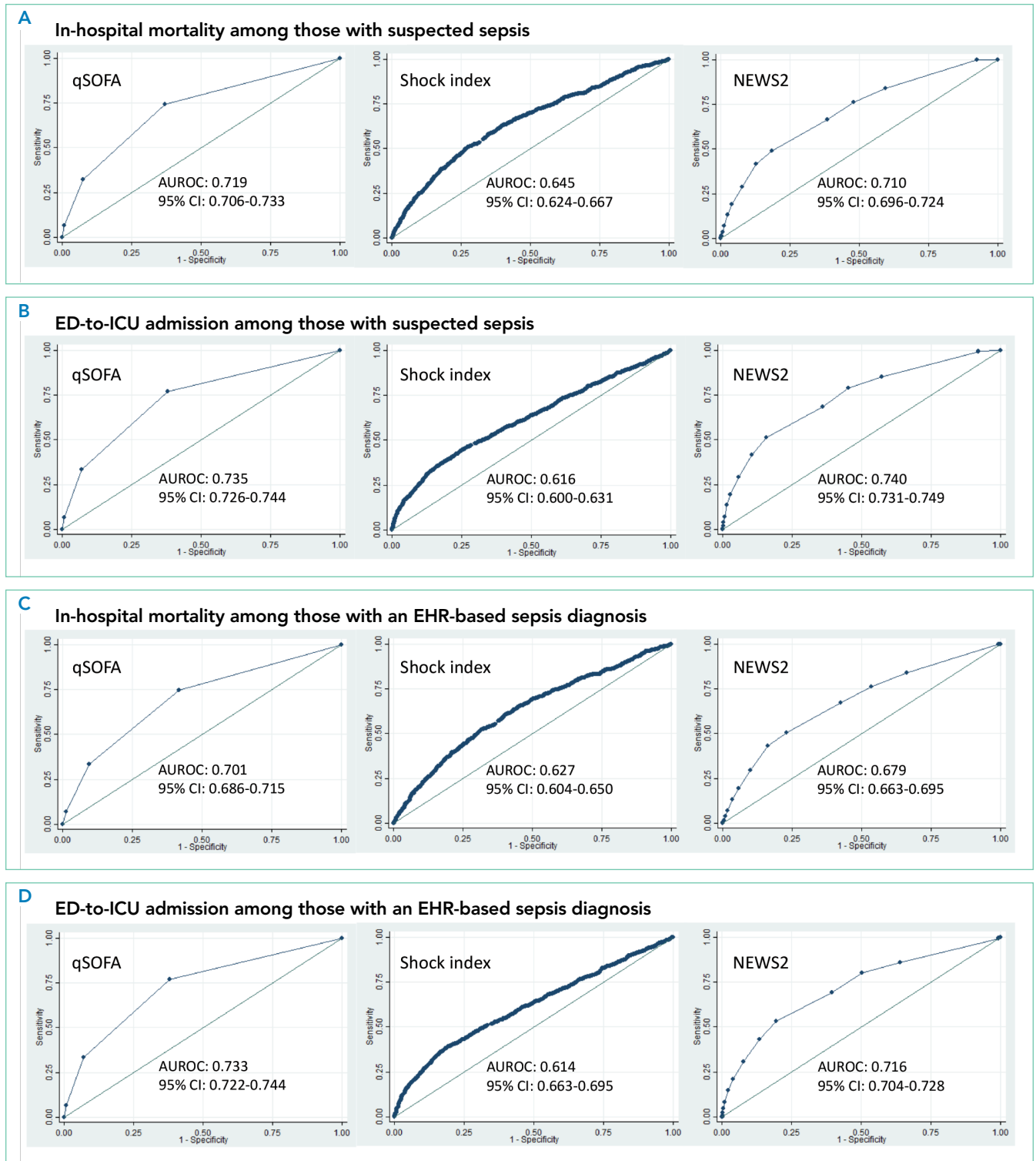


FIG. Receiver Operator Characteristic Curves for qSOFA, Shock Index, and NEWS2. The curves display the discrimination of each score to predict (A) in-hospital mortality among those with suspected sepsis, (B) emergency (ED)-to-intensive care unit (ICU) admission among those with suspected sepsis, (C) in-hospital mortality among those with electronic health record (EHR)-based sepsis diagnosis, and (D) ED-to-ICU admission among those with EHR-based sepsis diagnosis, considering the full range of possible scores measured in the cohort. The area under the receiver operator characteristic (AUROC) curves and 95% CIs are included. Abbreviations: NEWS2, National Early Warning System; qSOFA, quick Sequential Organ Failure Assessment.

These trends in sensitivity, specificity, and AUROC were consistent among those who met EHR criteria for a sepsis diagnosis. In the analysis of the three scoring systems using all available

cut-points, qSOFA and NEWS2 had the highest AUROCs, followed by the Shock Index.

Considering the rapid progression from organ dysfunction

to death in sepsis patients, as well as the difficulty establishing a sepsis diagnosis at triage,²³ providers must quickly identify patients at increased risk of poor outcomes when they present to the ED. Sepsis alerts often are built using SIRS criteria,²⁷ including the one used for sepsis surveillance at UCSF since 2012,²² but the white blood cell count criterion is subject to a laboratory lag and could lead to a delay in identification. Implementation of a point-of-care bedside score alert that uses readily available clinical data could allow providers to identify patients at greatest risk of poor outcomes immediately at ED presentation and triage, which motivated us to explore the predictive performance of qSOFA, the Shock Index, and NEWS2.

Our study is the first to provide a head-to-head comparison of the predictive performance of qSOFA, the Shock Index, and NEWS2, three easy-to-calculate bedside risk scores that use EHR data collected among patients with suspected sepsis. The Sepsis-3 guidelines recommend qSOFA to quickly identify non-ICU patients at greatest risk of poor outcomes because the measure exhibited predictive performance similar to the more extensive SOFA score outside the ICU.^{16,23} Although some studies have confirmed qSOFA's high predictive performance,²⁸⁻³¹ our test characteristics and AUROC findings are in line with other published analyses.^{4,6,10,17} The UK National Health Service is using NEWS2 to screen for patients at risk of poor outcomes from sepsis. Several analyses that assessed the predictive ability of NEWS have reported estimates in line with our findings.^{4,10,32} The Shock Index was introduced in 1967 and provided a metric to evaluate hemodynamic stability based on heart rate and systolic blood pressure.³³ The Shock Index has been studied in several contexts, including sepsis,³⁴ and studies show that a sustained Shock Index is associated with increased odds of vasopressor administration, higher prevalence of hyperlactatemia, and increased risk of poor outcomes in the ICU.^{13,14}

For our study, we were particularly interested in exploring how the Shock Index would compare with more frequently used severity scores such as qSOFA and NEWS2 among patients with suspected sepsis, given the simplicity of its calculation and the easy availability of required data. In our cohort of 23,837 patients, only 159 people had missing blood pressure and only 71 had omitted heart rate. In contrast, both qSOFA and NEWS2 include an assessment of level of consciousness that can be subject to variability in assessment methods and EHR documentation across institutions.¹¹ In our cohort, GCS within 30 minutes of ED presentation was missing in 72 patients, which could have led to incomplete calculation of qSOFA and NEWS2 if a missing value was not actually within normal limits.

Several investigations relate qSOFA to NEWS but few compare qSOFA with the newer NEWS2, and even fewer evaluate the Shock Index with any of these scores.^{10,11,18,29,35-37} In general, studies have shown that NEWS exhibits a higher AUROC for predicting mortality, sepsis with organ dysfunction, and ICU admission, often as a composite outcome.^{4,11,18,37,38} A handful of studies compare the Shock Index to SIRS; however, little has been done to compare the Shock Index to qSOFA or NEWS2,

scores that have been used specifically for sepsis and might be more predictive of poor outcomes than SIRS.³³ In our study, the Shock Index had a higher AUROC than either qSOFA or NEWS2 for predicting in-hospital mortality and ED-to-ICU admission measured as separate outcomes and as a composite outcome using standard cut-points for these scores.

When selecting a severity score to apply in an institution, it is important to carefully evaluate the score's test characteristics, in addition to considering the availability of reliable data. Tests with high sensitivity and NPV for the population being studied can be useful to rule out disease or risk of poor outcome, while tests with high specificity and PPV can be useful to rule in disease or risk of poor outcome.³⁹ When considering specificity, qSOFA's performance was superior to the Shock Index and NEWS2 in our study, but a small percentage of the population was identified using a cut-point of qSOFA ≥ 2 . If we used qSOFA and applied this standard cut-point at our institution, we could be confident that those identified were at increased risk, but we would miss a significant number of patients who would experience a poor outcome. When considering sensitivity, performance of NEWS2 was superior to qSOFA and the Shock Index in our study, but one-half of the population was identified using a cut-point of NEWS2 ≥ 5 . If we were to apply this standard NEWS2 cut-point at our institution, we would assume that one-half of our population was at risk, which might drive resource use towards patients who will not experience a poor outcome. Although none of the scores exhibited a robust AUROC measure, the Shock Index had the highest AUROC for in-hospital mortality and ED-to-ICU admission when using the standard binary cut-point, and its sensitivity and specificity is between that of qSOFA and NEWS2, potentially making it a score to use in settings where qSOFA and NEWS2 score components, such as altered mentation, are not reliably collected. Finally, our sensitivity analysis varying the binary cut-point of each score within our population demonstrated that the standard cut-points might not be as useful within a specific population and might need to be tailored for implementation, balancing sensitivity, specificity, PPV, and NPV to meet local priorities and ICU capacity.

Our study has limitations. It is a single-center, retrospective analysis, factors that could reduce generalizability. However, it does include a large and diverse patient population spanning several years. Missing GCS data could have affected the predictive ability of qSOFA and NEWS2 in our cohort. We could not reliably perform imputation of GCS because of the high missingness and therefore we assumed missing was normal, as was done in the Sepsis-3 derivation studies.¹⁶ Previous studies have attempted to impute GCS and have not observed improved performance of qSOFA to predict mortality.⁴⁰ Because manually collected variables such as GCS are less reliably documented in the EHR, there might be limitations in their use for triage risk scores.

Although the current analysis focused on the predictive performance of qSOFA, the Shock Index, and NEWS2 at triage, performance of these scores could affect the ED team's treatment decisions before handoff to the hospitalist team and the

expected level of care the patient will receive after in-patient admission. These tests also have the advantage of being easy to calculate at the bedside over time, which could provide an objective assessment of longitudinal predicted prognosis. Future work should assess the longitudinal performance of each of these scores among those with suspected sepsis and to determine the impact using these scores would have on clinical and resource utilization outcomes.

CONCLUSION

Local priorities should drive selection of a screening tool, balancing sensitivity, specificity, PPV, and NPV to achieve the institution's goals. qSOFA, Shock Index, and NEWS2 are risk stratification tools that can be easily implemented at ED triage using data available at the bedside. Although none of these scores performed strongly when comparing AUROCs, qSOFA was highly specific for identifying patients with poor outcomes, and NEWS2 was the most sensitive for ruling out those at high risk among patients with suspected sepsis. The Shock Index exhibited a sensitivity and specificity that fell between qSOFA and NEWS2 and also might be considered to identify those at increased risk, given its ease of implementation, particularly in settings where altered mentation is unreliably or inconsistently documented.

Acknowledgment

The authors thank the UCSF Division of Hospital Medicine Data Core for their assistance with data acquisition.

Disclosures: Dr Prasad reports personal fees from EpiExcellence, LLC, outside the submitted work. Dr Fang's institution received funding from the National Heart, Lung, and Blood Institute K24HL141354 and Patient-Centered Outcomes Research Institute. Dr Liu's institution received funding from the National Diabetes and Digestive and Kidney Diseases K24DK113381.

References

- Jones SL, Ashton CM, Kiehne LB, et al. Outcomes and resource use of sepsis-associated stays by presence on admission, severity, and hospital type. *Med Care*. 2016;54(3):303-310. <https://doi.org/10.1097/MLR.0000000000000481>
- Seymour CW, Gesten F, Prescott HC, et al. Time to treatment and mortality during mandated emergency care for sepsis. *N Engl J Med*. 2017;376(23):2235-2244. <https://doi.org/10.1056/NEJMoa1703058>
- Kumar A, Roberts D, Wood KE, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med*. 2006;34(6):1589-1596. <https://doi.org/10.1097/01.CCM.0000217961.75225.E9>
- Churpek MM, Snyder A, Sokol S, Pettit NN, Edelson DP. Investigating the impact of different suspicion of infection criteria on the accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores. *Crit Care Med*. 2017;45(11):1805-1812. <https://doi.org/10.1097/CCM.0000000000002648>
- Abdullah SMOB, Sørensen RH, Dessau RBC, Sattar SMRU, Wiese L, Nielsen FE. Prognostic accuracy of qSOFA in predicting 28-day mortality among infected patients in an emergency department: a prospective validation study. *Emerg Med J*. 2019;36(12):722-728. <https://doi.org/10.1136/emered-2019-208456>
- Kim KS, Suh GJ, Kim K, et al. Quick Sepsis-related Organ Failure Assessment score is not sensitive enough to predict 28-day mortality in emergency department patients with sepsis: a retrospective review. *Clin Exp Emerg Med*. 2019;6(1):77-83. <https://doi.org/10.15441/ceem.17.294>
- National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. Royal College of Physicians; 2017.
- Brink A, Alisma J, Verdonschot RJCG, et al. Predicting mortality in patients with suspected sepsis at the emergency department: a retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. *PLoS One*. 2019;14(1):e0211133. <https://doi.org/10.1371/journal.pone.0211133>
- Redfern OC, Smith GB, Prytherch DR, Meredith P, Inada-Kim M, Schmidt PE. A comparison of the Quick Sequential (Sepsis-Related) Organ Failure Assessment Score and the National Early Warning Score in non-ICU patients with/without infection. *Crit Care Med*. 2018;46(12):1923-1933. <https://doi.org/10.1097/CCM.0000000000003359>
- 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(7):906-911. <https://doi.org/10.1164/rccm.201604-0854OC>
- Goulden R, Hoyle MC, Monis J, et al. qSOFA, SIRS and NEWS for predicting in-hospital mortality and ICU admission in emergency admissions treated as sepsis. *Emerg Med J*. 2018;35(6):345-349. <https://doi.org/10.1136/emered-2017-207120>
- Biney I, Shepherd A, Thomas J, Mehari A. Shock Index and outcomes in patients admitted to the ICU with sepsis. *Chest*. 2015;148(suppl 4):337A. <https://doi.org/https://doi.org/10.1378/chest.2281151>
- Wira CR, Francis MW, Bhat S, Ehrman R, Conner D, Siegel M. The shock index as a predictor of vasopressor use in emergency department patients with severe sepsis. *West J Emerg Med*. 2014;15(1):60-66. <https://doi.org/10.5811/westjem.2013.7.18472>
- Berger T, Green J, Horeczko T, et al. Shock index and early recognition of sepsis in the emergency department: pilot study. *West J Emerg Med*. 2013;14(2):168-174. <https://doi.org/10.5811/westjem.2012.8.11546>
- Middleton DJ, Smith TO, Bedford R, Neilly M, Myint PK. Shock Index predicts outcome in patients with suspected sepsis or community-acquired pneumonia: a systematic review. *J Clin Med*. 2019;8(8):1144. <https://doi.org/10.3390/jcm8081144>
- 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(8):762-774. <https://doi.org/10.1001/jama.2016.0288>
- Abdullah S, Sørensen RH, Dessau RBC, Sattar S, Wiese L, Nielsen FE. Prognostic accuracy of qSOFA in predicting 28-day mortality among infected patients in an emergency department: a prospective validation study. *Emerg Med J*. 2019;36(12):722-728. <https://doi.org/10.1136/emered-2019-208456>
- Usman OA, Usman AA, Ward MA. Comparison of SIRS, qSOFA, and NEWS for the early identification of sepsis in the Emergency Department. *Am J Emerg Med*. 2018;37(8):1490-1497. <https://doi.org/10.1016/j.ajem.2018.10.058>
- Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27. <https://doi.org/10.1097/00005650-199801000-00004>
- van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care*. 2009;47(6):626-633. <https://doi.org/10.1097/MLR.0b013e31819432e5>
- Prin M, Wunsch H. The role of stepdown beds in hospital care. *Am J Respir Crit Care Med*. 2014;190(11):1210-1216. <https://doi.org/10.1164/rccm.201406-1117PP>
- Narayanan N, Gross AK, Pintens M, Fee C, MacDougall C. Effect of an electronic medical record alert for severe sepsis among ED patients. *Am J Emerg Med*. 2016;34(2):185-188. <https://doi.org/10.1016/j.ajem.2015.10.005>
- Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*. 2016;315(8):801-810. <https://doi.org/10.1001/jama.2016.0287>
- Rhee C, Dantes R, Epstein L, et al. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. *JAMA*. 2017;318(13):1241-1249. <https://doi.org/10.1001/jama.2017.13836>
- Safari S, Baratloo A, Elfil M, Negida A. Evidence based emergency medicine; part 5 receiver operating curve and area under the curve. *Emerg (Tehran)*. 2016;4(2):111-113.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.
- Kangas C, Iverson L, Pierce D. Sepsis screening: combining Early Warning Scores and SIRS Criteria. *Clin Nurs Res*. 2021;30(1):42-49. <https://doi.org/10.1177/1054773818823334>
- 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(3):301-308.

<https://doi.org/10.1001/jama.2016.20329>

29. Finkelsztain EJ, Jones DS, Ma KC, et al. Comparison of qSOFA and SIRS for predicting adverse outcomes of patients with suspicion of sepsis outside the intensive care unit. *Crit Care*. 2017;21(1):73. <https://doi.org/10.1186/s13054-017-1658-5>
30. Canet E, Taylor DM, Khor R, Krishnan V, Bellomo R. qSOFA as predictor of mortality and prolonged ICU admission in Emergency Department patients with suspected infection. *J Crit Care*. 2018;48:118-123. <https://doi.org/10.1016/j.jcrc.2018.08.022>
31. Anand V, Zhang Z, Kadri SS, Klompas M, Rhee C; CDC Prevention Epicenters Program. Epidemiology of Quick Sequential Organ Failure Assessment criteria in undifferentiated patients and association with suspected infection and sepsis. *Chest*. 2019;156(2):289-297. <https://doi.org/10.1016/j.chest.2019.03.032>
32. Hamilton F, Arnold D, Baird A, Albur M, Whiting P. Early Warning Scores do not accurately predict mortality in sepsis: A meta-analysis and systematic review of the literature. *J Infect*. 2018;76(3):241-248. <https://doi.org/10.1016/j.jinf.2018.01.002>
33. Koch E, Lovett S, Nghiem T, Riggs RA, Rech MA. Shock Index in the emergency department: utility and limitations. *Open Access Emerg Med*. 2019;11:179-199. <https://doi.org/10.2147/OAEM.S178358>
34. Yussof SJ, Zakaria MI, Mohamed FL, Bujang MA, Lakshmanan S, Asaari AH. Value of Shock Index in prognosticating the short-term outcome of death for patients presenting with severe sepsis and septic shock in the emergency department. *Med J Malaysia*. 2012;67(4):406-411.
35. Siddiqui S, Chua M, Kumaresh V, Choo R. A comparison of pre ICU admission SIRS, EWS and qSOFA scores for predicting mortality and length of stay in ICU. *J Crit Care*. 2017;41:191-193. <https://doi.org/10.1016/j.jcrc.2017.05.017>
36. Costa RT, Nassar AP, Caruso P. Accuracy of SOFA, qSOFA, and SIRS scores for mortality in cancer patients admitted to an intensive care unit with suspected infection. *J Crit Care*. 2018;45:52-57. <https://doi.org/10.1016/j.jcrc.2017.12.024>
37. Mellhammar L, Linder A, Tverring J, et al. NEWS2 is Superior to qSOFA in detecting sepsis with organ dysfunction in the emergency department. *J Clin Med*. 2019;8(8):1128. <https://doi.org/10.3390/jcm8081128>
38. Szakmany T, Pugh R, Kopczynska M, et al. Defining sepsis on the wards: results of a multi-centre point-prevalence study comparing two sepsis definitions. *Anaesthesia*. 2018;73(2):195-204. <https://doi.org/10.1111/anae.14062>
39. Newman TB, Kohn MA. *Evidence-Based Diagnosis: An Introduction to Clinical Epidemiology*. Cambridge University Press; 2009.
40. Askim Å, 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(1):56. <https://doi.org/10.1186/s13049-017-0399-4>

Journal of Hospital Medicine®

Leading By Example

journalofhospitalmedicine.com

#HowWeHospitalist

