OREGON HEALTH AND SCIENCE UNIVERSITY
OFFICE OF CLINICAL INTEGRATION AND EVIDENCE-BASED PRACTICE

Evidence-Based Practice Summary

Tools for Identification of Sepsis in Hospitalized and Emergency Department Patients

Prepared for: Mortality A3 Management Guidance Team
Authors: Tovah Kohl, MA; Elizabeth Crabtree, PhD, MPH

BACKGROUND AND RATIONALE

Sepsis is the most expensive condition treated in the hospital, resulting in an aggregate cost of $20.3 billion or 5.2% of total cost for all hospitalizations in the United States. Rates of sepsis and sepsis related mortality are rising in the United States. Timely treatment and identification of sepsis decreases morbidity, mortality, and costs (Makam 2015).

In order to improve outcomes, there have been a variety of tools developed to identify sepsis and predict mortality. These tools use a number of different combinations of culture orders and antibiotic use that likely represent different degrees of suspicion of infection. However, due to a lack of a gold standard for the definition or identification of sepsis, the criteria used for infection suspicion are subjective and can only reflect professional preference and not objective evidence of infection. These various combinations may result in important differences in patient outcomes, and accuracy of tools designed to risk stratify septic patients (Churpek 2017).

The purpose of this evidence brief is to search for and appraise literature on the various tools used in the hospital and ED settings to identify sepsis. Tools included in the analysis are well-known, validated instruments, and the brief excludes homegrown tools or modified versions of validated instruments, created for specific hospital settings. The evidence brief does not address implementation strategies, nor does it assess the impact of using various tools on patient-important outcomes.

ASK THE QUESTION

For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?

SEARCH FOR EVIDENCE

Databases included Ovid MEDLINE, and National Guideline Clearinghouse,

Ovid MEDLINE search strategy included:

1 exp Sepsis/ (117761)
2 exp Mass Screening/ (122223)
3 exp Hospitalization/ (217535)
4 1 and 2 and 3 (12)
5 qsofa.mp. (46)
6 q-sofa.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (1)
7 quick sofa.mp. (11)
8 5 or 7 (47)
9 exp Systemic Inflammatory Response Syndrome/ (121646)
10 exp Sepsis/di, ep, sn, pc [Diagnosis, Prevention & Control] (33672)
11 exp "Sensitivity and Specificity"/ (558507)
12 10 and 11 (3035)
13 limit 12 to comparative study (704)
14 exp hospitals/ (259996)
15 exp Hospital Administration/ (246325)
16 3 or 14 or 15 (639179)
17 12 and 16 (363)
18 exp Biomarkers/ (878118)
19 10 and 16 and 18 (211)
20 12 and 18 (758)
21 16 and 20 (117)
22 19 or 21 (211)
23 ((sirs or Systemic Inflammatory Response Syndrome or qsofa or q-sofa or quick sofa or quicksofa or (SEPSIS adj2 (ORGAN* adj FAIL*) adj3 ASSESSMENT) or mews or modified early warning system*) adj7 (predict* or recogniz* or identif* or detect* or diagnos* or prognos* or death* or die or dying or died or dies or mortal* or find or found or accura* or inaccura*) adj10 (sepsis* or septic*)).mp. (406)
24 4 or 17 or 19 or 21 or 23 (849)
25 limit 24 to english language (784)
26 remove duplicates from 25 (701)
27 limit 26 to (comparative study or evaluation studies or validation studies or controlled clinical trial or guideline or meta analysis or randomized controlled trial or systematic reviews) (222)
28 exp Epidemiologic Studies/ (2301239)
29 26 and 28 (420)
30 29 not 27 (289)
31 26 not (27 or 29) (190)

Filters/limits included research articles published in English in the last 5 years

CRITICALLY ANALYZE THE EVIDENCE

The literature search resulted in 511 studies reporting on the test performance characteristics of various sepsis tools and biomarkers. Due to this fact, we limited the literature to 17 studies published in the last five years focusing on well-known and validated tools. In order to simplify the review process, the evidence appraisal tables have been grouped by scoring tool. These categories include: (1) qSOFA; (2) qSOFA vs SIRS; (3) Australasian Triage Scale; (4) SIRS; (5) MEWS; (6) SOFA vs MEDS; (7) APACHE II; (8) Step-by-Step; (9) NEWS; and (10) MISSED.

1. **Tool: qSOFA:** Four diagnostics studies evaluated the test characteristics of the quick Sepsis Related Organ Failure Assessment (qSOFA) tool. The first study (Chen 2016) investigated the predictive performance of the qSOFA tool for mortality, hospitalization, and ICU admissions in patients presenting to the ED with pneumonia. The study found that the sensitivity of qSOFA for mortality, hospitalization, and ICU admissions with a cut-off value of 2, respectively, was 12%, 8%, 18% and specificity was 97%, 97%, and 96%. The second study (Churpek 2017) attempted to determine the impact of different criteria on the accuracy of screening tools on patients hospitalized outside of the ICU at an academic medical center. They found that for identifying potentially infected patients the median area-under-the- curve (AUC) of 0.65 and the sensitivity of qSOFA >1 ranged from 69% to 73% at 51-57% specificity. For qSOFA $\geq$2, the sensitivity ranged from 18% to 23% at 91-94% specificity. The third study (Wang 2016) investigated the performance of qSOFA in predicting mortality and ICU admission in patients with clinically diagnosed infection in the ED. In patients with qSOFA scores < 2 and $\geq$ 2, 28-day mortality rates were 17.4% and 42.9% (P< 0.001), and ICU admission rates were 16.0% and 33.3% (P< 0.001). The fourth study (Kim 2017) evaluated the predictive performance of the qSOFA score as a screening tool for sepsis, mortality, and ICU admission in patients with febrile neutropenia (FN). The qSOFA score was a significant factor in predicting sepsis [OR 1.676, 95% CI (1.014-2.772)] and ICU admission [OR 2.350, 95% CI (1.273-4.338)]. The qSOFA tool showed a low sensitivity (0.14, 0.2, and 0.23) but high specificity (0.98, 0.97, and 0.97) in predicting sepsis, 28-day mortality, and ICU admissions.

   Quality of Evidence: Low, AUC for identifying sepsis 0.65, sensitivity for predicting mortality range 2%-49% and specificity 97%, sensitivity for identifying sepsis range 14%-73% and specificity range 51%-98%

2. **Tool: qSOFA vs SIRS:** Four diagnostic studies compared the test characteristics of the qSOFA tool and the Systematic Inflammatory Response Syndrome (SIRS) tool. One study (Moskowitz 2017) investigated the related predictive performance of both the qSOFA and the SIRS criteria in patients with suspected infection in the ED. The AUC when using qSOFA to predict critical care intervention (CCI) was 0.74 (95% CI, 0.73-0.74) and 0.71 (0.69-0.72) when used to predict in-hospital mortality. The AUC of SIRS to predict CCI was 0.68 (0.68-0.69) and, when used to predict in-hospital mortality, 0.66 (0.65-0.68). The AUC of qSOFA to predict CCI was significantly higher than that for SIRS ($p < 0.001$). The second study (Askim 2017) evaluated the clinical usefulness of qSOFA as a risk stratification tool for patients admitted with infection compared to traditional SIRS

criteria. When comparing the ability to correctly identify patients with severe sepsis, qSOFA ≥2 was able to identify 30.6% of all cases versus SIRS ≥2 was able to identify 74.1% of all cases. The sensitivity of qSOFA ≥2 was 0.32 and SIRS ≥2 was 0.74 and the specificity was 0.98 and 0.72, respectively. The third study (Williams 2017) compared the diagnostic accuracy of SIRS and qSOFA for organ dysfunction with the updated Sepsis-3 definition of organ dysfunction. The sensitivity of qSOFA ≥2 was 0.297 and 0.721 for SIRS, while the AUC was 0.73 for qSOFA and 0.72 for SIRS. The fourth study (Freund 2017) prospectively validated the qSOFA score and the SIRS score to predict mortality. The AUC for qSOFA was 0.80 compared to 0.65 for SIRS to predict in-hospital mortality.

Quality of Evidence: Low to support performance of qSOFA vs. SIRS; AUC qSOFA to predict mortality range 0.71-0.80. AUC qSOFA to identify sepsis 0.73, sensitivity range 29.7%-30.6%, specificity 98%. AUC SIRS to predict mortality range 0.65-0.66. AUC SIRS to predict sepsis 0.72, sensitivity range 72.1%-74%, specificity 72%.

3. **Tool: Australasian Triage Scale:** One study evaluated the test characteristics of the Australasian Triage Scale (ATS). It investigated the accuracy and the validity of the ATS to identify and manage patients with severe sepsis in a timely manner. The overall sensitivity of the ATS to identify severe sepsis was 71% (95% CI 0.676-0.741) with a PPV of 0.586 (95% CI 0.554-0.559) (Chamberlain, Willis, Clark, & Brideson, 2015).

   Quality of Evidence: Low; sensitivity 71% and PPV 0.586 for identifying sepsis

4. **Tool: SIRS:** Two diagnostic studies evaluated the test characteristics of the SIRS tool. The first study (Churpek 2017) attempted to determine the impact of different criteria on the accuracy of screening tools on patients hospitalized outside of the ICU at an academic medical center. Accuracy for predicting composite outcome (death+ ICU admissions) across all criteria of infection for SIRS was AUC 0.60 (95% CI 0.60-0.61) and the sensitivity of SIRS≥2 ranged from 66%-69%. Maguire (2016) examined, in the setting of maternal bacteremia, the implications for the diagnosis of maternal sepsis of customizing the SIRS criteria for physiologic changes of pregnancy. Of the 93 women in the study with bacteremia, 61 (66%) would have been diagnosed with sepsis based on standard SIRS compared with 52 (56%) based on customized SIRS (P=0.18).

   Quality of Evidence: Low; AUC for predicting mortality 0.60, sensitivity ranged from 56%-69% for correctly identifying sepsis and predicting mortality

5. **Tool: MEWS:** Two diagnostic studies were found researching the test characteristics of the modified early warning score (MEWS). The first study (Churpek 2017) attempted to determine the impact of different criteria on the accuracy of screening tools on patients hospitalized outside of the ICU at an academic medical center. Accuracy for predicting composite outcome (death+ ICU admissions) across all criteria of infection for MEWS was AUC 0.65. The specificity ranged from 51-57% compared to qSOFA>1 cutoff and the sensitivity ranged from 17-21% at the qSOFA>2 cutoff. The second study (Cildir 2013) evaluated the ability of the modified early warning score (MEWS) to predict prognosis in patients who are diagnosed with sepsis. The ability for MEWS to correctly predict 28-day mortality in patients with sepsis was 87.5% (30.4% specificity and an AUC of 0.574) and 48.5% (67.0% specificity and an AUC of 0.596) for patients with severe sepsis.

Quality of Evidence: Low; AUC for predicting mortality range of 0.574-0.65, specificity range 30.4%-67%, sensitivity range 17%-87.5%

6. **Tool: SOFA vs MEDS:** Three diagnostic studies evaluated the test characteristics of the Sepsis Related Organ Failure Assessment (SOFA) tool to the mortality in emergency department sepsis (MEDS) tool. The first (MacDonald 2014) sought to predict mortality in the ED in patients with features suggesting severe sepsis or septic shock using MEDS and SOFA scores. They found that pairwise, comparisons of the AUC MEDS (0.81) versus SOFA (0.78) was not statistically significant (p = 0.37). The second study (Wang 2016) investigated the performance of qSOFA in predicting mortality and intensive care unit (ICU) admission in patients with clinically diagnosed infection and to compare its performance with that of MEDS, APACHE II, and SOFA. The AUC of MEDS (0.751) is similar to that of SOFA (0.729) and APACHE II (0.732) in predicting 28-day mortality. The AUC of qSOFA, SOFA, MEDS, and APACHE II in predicting ICU admission were 0.636, 0.682, 0.661, and 0.640, respectively. They found no significant differences among the score systems. The third study (Williams 2016) compared the performance of scores in patient subgroups with increasing mortality: infection without systemic inflammatory response syndrome, sepsis, severe sepsis, and septic shock. AUC values for the entire cohort were; MEDS score of 0.92 and SOFA score of 0.86. All scores overestimated mortality, but closest concordance between predicted and observed mortality was seen with Mortality in Emergency Department Sepsis (MEDS) score.
   Level of Evidence: Moderate to support no difference in performance between SOFA and MEDS; AUC of MEDS for predicting mortality range of 0.751-0.92 and the AUC of SOFA for predicting mortality range of 0.729-0.86

7. **Tool: APACHE II:** Two diagnostic studies assessed the accuracy of using the Acute Physiology and Chronic Health Evaluation (APACHE) II in predicting mortality. Both studies included patients presenting to the ED and admitted with a presumed infection. The first, a retrospective cohort study of 477 patients, compared the performance of qSOFA to a number of other tools, including the APACHE II for predicting mortality and ICU admissions. It found no significant differences between scoring systems. The AUC for APACHE II in predicting 28-day mortality was 0.732, similar to that of SOFA (0.729) and qSOFA (0.666), and MEDS (0.751). The AUC of qSOFA, SOFA, MEDS, and APACHE II in predicting ICU admission were 0.636, 0.682, 0.661, and 0.640, respectively (Wang, Chen, Guo, Mei, & Yang, 2016). A large prospective observational study of 8871 patients, compared a number of severity of illness scores in predicting mortality. The AUROC for the Mortality in Emergency Department Sepsis score was 0.92, the Simplified APACHE II and APACHE II scores were 0.90**,** the Sequential Organ Failure Assessment score was 0.86, and the Severe Sepsis Score was 0.82 (Williams, Greenslade, Chu, Brown, & Lipman, 2016).
   Quality of Evidence: Moderate; AUC for predicting mortality range of 0.732 – 0.90

8. **Tool: Step-by-Step:** One study was identified specific to pediatric patients. A prospective observational study of 2185 infants less than or equal to 90 days, presenting to an ED with fever without source, compared the performance of different scoring systems in identifying patients at low risk of IBI (isolation of bacterial pathogen in a blood or cerebrospinal fluid). The purpose of the study was to validate the Step-by-Step approach and compare it

with the Rochester criteria and the Lab-score. Sensitivity and negative predictive value for ruling out an IBI were 92.0% and 99.3% for the Step by Step, 81.6% and 98.3% for the Rochester criteria, and 59.8% and 98.1% for the Lab-score. Seven infants with an IBI were misclassified by the Step by Step, 16 by Rochester criteria, and 35 by the Lab- score (Gomez et al., 2016).

Quality of Evidence: Low; Sensitivity and NPV for ruling out an IBI 92% and 99.3%, respectively

9. **Tool: NEWS:** One retrospective observational study identified evaluated the relationship between the initial national EWS (NEWS) in the ED and the diagnosis of severe sepsis and septic shock in 500 non-trauma adult patients presenting to the ED with a Manchester Triage System category 1 – 3. The AUC for NEWS to identify patients at risk for severe sepsis or septic shock was 0.89 (95% CI 0.84 to 0.94). A NEWS of 3 or more at ED triage had a sensitivity of 92.6% (95% CI 74.2% to 98.7%) and a specificity of 77% (95% CI 72.8% to 80.6%) to severe sepsis or septic shock at ED triage (Keep et al., 2016).

   Quality of Evidence: Low; AUC in identifying patients at risk for severe sepsis or septic shock 0.89

10. **Tool: MISSED:** A retrospective observational study of 674 ED patients admitted with a diagnosis of sepsis sought to validate the simplified MISSED score for predicting all-cause mortality in the ED population. The AUC for predicting mortality for the simplified MISSED score was 0.74 [95% confidence interval (CI) 0.70- 0.77; P< 0.00011]. The odds ratio for mortality for a score 2 or more was 5.01 (95% CI 2.93-8.57; P < 0.0001), and that for ICU admission was 3.00 (95% CI 1.70-5.28; P= 0.0001) (Sivayoham, Holmes, Cecconi, & Rhodes, 2015).

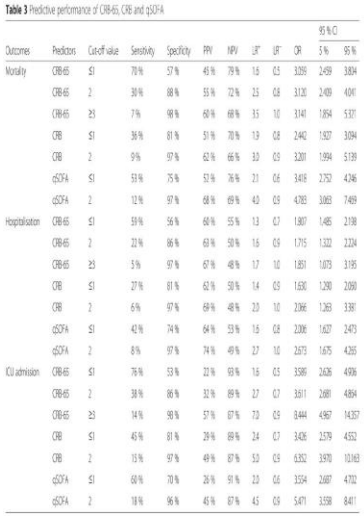
    Quality of Evidence: Low; AUC for predicting mortality 0.74

ACCURACY OF DIFFERENT TOOLS FOR PREDICTING MORTALITY

Algorithm	AUC
qSOFA vs SIRS	qSOFA 0.71-0.80 vs SIRS 0.65-0.66
SIRS	0.60
MEWS	0.574-0.65
SOFA vs MEDS	MEDS 0.751-0.92 vs. SOFA 0.729-0.86
APACHE II	0.732 – 0.90
MISSED	0.74

Overall, the evidence for the test characteristics for the various tools was rated as low, with a few comparisons rated as moderate. There were limited studies found reporting on the accuracy of the tools to predict sepsis in patients. There were issues with inconsistency and indirectness with most studies evaluating the prediction of mortality or ICU admissions rather than the prediction of sepsis among varying patient populations. We cannot conclude with great certainty which tool is most effective in identifying sepsis early in patients.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if:
Tool: qSOFA						☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)* ☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)* ☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness*
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	
Total # of Studies: 4 # of Diagnostic Studies: 4						
Chen, Y. X., et al. (2016). *Critical Care*	To investigate the predictive performance of qSOFA, CRB-65 (confusion, respiratory rate >=30/minute, systolic blood pressure <90 mmHg or diastolic blood pressure <=60 mmHg, age >=65 years) and CRB (confusion, respiratory rate >=30/minute, systolic blood pressure <90 mmHg or diastolic blood pressure <=60 mmHg) for mortality, hospitalisation	Retrospective cohort study.	1641 **adult patients presenting with pneumonia.**	**Compared with patients with qSOFA scores <2, those with qSOFA scores ≥2 had a one- to six-fold increase in the prevalence of hospitalization.** In baseline risk classes 2-4, the fold change of hospitalization was two- to six-fold for qSOFA and was more obvious than that for CRB/CRB-65 (one- to four-fold). In baseline risk class ≥5. Fold changes of the three systems were similar. **Patients with qSOFA scores ≥2 had a one- to three-fold increase in the prevalence of ICU admission compared with patients with qSOFA scores <2.** **Twenty-eight-day mortality and ICU admission of patients with qSOFA scores of 2 and 3 were much higher than those of patients with the same CRB-65 scores (P < 0.01).** **Patients with qSOFA scores of 0, 1, 2 and 3 were associated with, respectively, mortality of 16.3 %, 24.4 %, 48.2 % and 68.4 %; prevalence of hospitalization of**	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

	and ICU admission in patients with pneumonia in the emergency department (ED)..			**37.2 %, 47.4 %, 61.6 % and 73.7 %; and prevalence of ICU admission of 9.3 %, 9.1 %, 22.4 % and 45.3 %.** **Table 3** Predictive performance of CRB-65, CRB and qSOFA 		*of drug, only small, positive studies found)* <u>Increase Quality Rating if:</u> ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low
Churpek, M. M., et al. (2017). *Critical Care Medicine*	To determine the impact of differing criteria on the accuracy of sepsis screening tools and early warning scores.	Prospective cohort study of all **patients hospitalized outside of the ICU at an academic medical center**.	150,288 patients were hospitalized and 53,849 met at least one infection criteria.	***For identifying potentially infected patients:*** **Median AUC, 0.65 (range, 0.62-0.66)** **The sensitivity of qSOFA >1 ranged from 69% to 73% at 51-57% specificity.** **The sensitivity of qSOFA <u>≥</u>2, which ranged from a sensitivity of 18% to 23% at 91-94% specificity.**	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

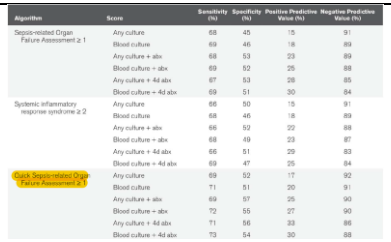
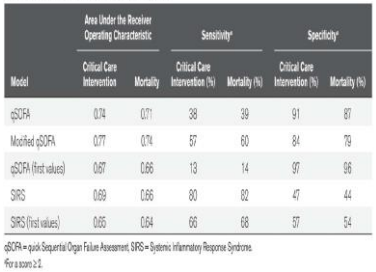
				 TABLE 3. Accuracy of the Different Scores Across Criteria for Suspicion of Infection, With Cut-Offs Matched to the Sensitivity of Quick Sepsis-Related Organ Failure Assessment ≥ 2 		
Wang, J. Y., et al. (2016). *American Journal of Emergency Medicine*	The objectives of this study are to investigate the performance of qSOFA in predicting mortality and intensive care unit (ICU) admission in patients with clinically diagnosed infection and to compare its performance with that of MEDS, APACHE II, and SOFA.	Retrospective cohort study.	477 **patients clinically diagnosed with infection in the ED.**	**All scores were higher in nonsurvivors and ICU patients than in survivors and non-ICU patients (P< 0.001).** **In patients with qSOFA scores less than 2 and greater than or equal to 2, 28-day mortality rates were 17.4% and 42.9% (P< 0.001), and ICU admission rates were 16.0% and 33.3% (P< 0.001).**	Study Limitations = ☐ None ***Diagnostic Studies*** ☒ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☐ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

				Table 3 Performance of qSOFA in predicting mortality and ICU admission Outcomes Cutoff Sensitivity Specificity PPV NPV LR+ LR− OR 95% CI 28-d mortality 2 42.9% 82.6% 61.6% 69.8% 2.47 0.69 3.6 2.3 5.4 ICU admission 2 33.3% 84.6% 57.8% 65.6% 2.08 0.79 2.6 1.7 4.1 Abbreviations: PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR−, negative likelihood ratio; OR, odd ratio.		
Kim, M., et al. (2017). *Supportive Care in Cancer*	To evaluate the predictive performance of the qSOFA score as a screening tool for sepsis, mortality, and intensive care unit (ICU) admission in patients with febrile neutropenia (FN).	Prospective cohort study. The qSOFA and SIRS scores were calculated retrospectively using the preexisting data. The primary outcome was the development of sepsis. The secondary outcomes were ICU admission and 28-day mortality.	**615 patients with febrile neutropenia**	the qSOFA score was a significant factor in predicting **sepsis [OR 1.676,** 95% CI (1.014-2.772)] **and ICU admission [OR 2.350**, 95% CI (1.273-4.338)] **qSOFA showed a low sensitivity (0.14, 0.2, and 0.23) but high specificity (0.98, 0.97, and 0.97) in predicting sepsis, 28-day mortality, and ICU admission.**	Study Limitations = ☐ None ***Diagnostic Studies*** ☒ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☐ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?

Tool: qSOFA vs SIRS

Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*
Total # of Studies: 4 # of Systematic Reviews: 0 # of RCTs: 0 # of Non-Randomized Studies: 0 # of Diagnostic Studies: 4					
Moskowitz, A., et al. (2017). *Critical Care Medicine*	To investigate the related predictive performance of both the quick Sequential Organ Failure Assessment and the Systemic Inflammatory Response Syndrome criteria.	Retrospective cohort study.	24,164 **patients with suspected infection who presented to the emergency department** and were admitted to the hospital between January 2010 and December 2014	**The AUROC when using qSOFA to predict Critical Care Intervention (CCI) was 0.74 (95% CI, 0.73-0.74) and 0.71 (0.69-0.72) when used to predict inhospital mortality.** **The AUROC of SIRS to predict CCI was 0.68 (0.68-0.69) and, when used to predict mortality, 0.66 (0.65-0.68).** **The AUROC of qSOFA to predict receipt of CCI was significantly higher than that for SIRS (p < 0.001)** **TABLE 2. Model Diagnostic Characteristics**  qSOFA = quick Sequential Organ Failure Assessment, SIRS = Systemic Inflammatory Response Syndrome. *For a score ≥ 2.	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☐ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis
Askim, A., et al. (2017). *Scandinavian Journal of Trauma,*	To evaluate the clinical usefulness of qSOFA as a risk stratification	Prospective cohort study. At arrival in the ED, vital signs were recorded and all patients were triaged according to RETTS vital signs, presenting infection, and sepsis	1535 admitted patients	Of the 1535 patients, 26 (1.7%) died within 7 days and only four (15.4%) of them were identified by the qSOFA ≥ 2 in the ED compared to 17 (65.4%) for SIRS ≥2	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner

Lower Quality Rating if:

☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)*

☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)*

☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)*

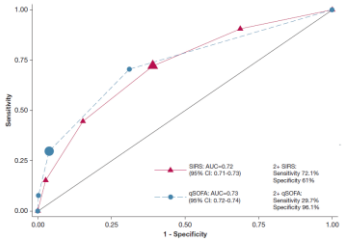
☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)*

Increase Quality Rating if:

☐ Large Effect

☐ Dose-response gradient

☐ Plausible confounders or other biases increase certainty of effect

Resuscitation & Emergency Medicine	tool for **patients admitted with infection** compared to traditional SIRS criteria or our triage system; the Rapid Emergency Triage and Treatment System (RETTS).	symptoms. These admission data were also used to calculate qSOFA and SIRS.		***SIRS >2 ability to identify severe sepsis:*** **74.1% of all cases (80/108) sensitivity 0.74 (95% CI, 0.65-0.82) specificity 0.72 (95%CI, 0.70-0.75) PPV 0.18 (95%CI 0.16-0.19) NPV 0.97 (95% CI 0.96-0.98)** ***qSOFA > 2 ability to identify severe sepsis*: 30.6% of all cases (33/103) sensitivity: 0.32 (95% CI 0.23-0.42) specificity: 0.98 (95% CI 0.97-0.99) PPV 0.57 (95% CI 0.45-.0.68) NPV 0.95 (95% CI 0.94-0.96)**	☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☐ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☒ Failure to include all patients in analysis	Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low
Williams, J. M., et al. (2017). *Chest*	To determine the prognostic impact of SIRS, compare the diagnostic accuracy of SIRS and qSOFA for organ dysfunction, and (compare standard (Sepsis-2) and revised (Sepsis-3) definitions for organ dysfunction in ED patients with infection.	Prospective cohort study **of patients admitted to the ED over 3 years with presumed infection.** Data was collected in order to calculate SIRS, qSOFA, SOFA, comorbidity, and mortality	8,871 patients	***Prediction of Sepsis-3 organ dysfunction:*** **SIRS: AUC=0.72 (95% CI 0.71-0.73) 2+SIRS score sensitivity: 72.1%(95% CI 70.3-74.1) specificity 61% (95% CI, 60.0-62.3)** **qSOFA: AUC 0.73 (95% CI 0/72-0.74) 2+qSOFA score sensitivity: 29.7% (95% CI, 27.9-31.8) specificity 96.1% (95%CI 95.7-96.6)** 	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☐ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	
Freund, Y., et al. (2017). *JAMA*	To prospectively validate qSOFA as a mortality predictor and	Prospective cohort study of **patients who visited the ED with suspected infection**	878 patients	**qSOFA to predict in-hospital mortality AUC 0.80** (95% CI, 0.74-0.85) sensitivity 70% (95% CI 59-80) specificity 79% (95%CI 76-82)	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner	

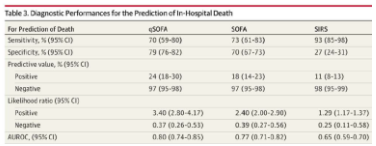
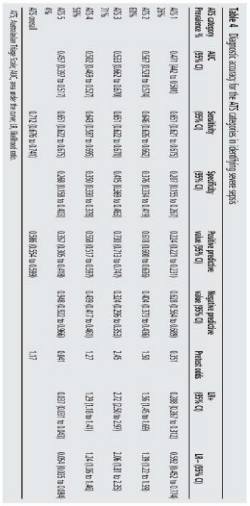
	compare the performances of the new sepsis criteria to the previous ones.			**SIRS to predict in-hospital mortality AUC 0.65** (95% CI, 0.59-0.70) sensitivity 93% (95% CI 85-98) Specificity 27% (95%CI 24-31) 	☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

Table 3. Diagnostic Performances for the Prediction of In-Hospital Death

For Prediction of Death	qSOFA	SOFA	SIRS
Sensitivity, % (95% CI)	70 (59-80)	73 (61-83)	93 (85-98)
Specificity, % (95% CI)	79 (76-82)	70 (67-73)	27 (24-31)
Predictive value, % (95% CI)			
Positive	24 (18-30)	18 (14-23)	11 (8-13)
Negative	97 (95-98)	97 (95-98)	98 (95-99)
Likelihood ratio (95% CI)			
Positive	3.40 (2.80-4.17)	2.40 (2.00-2.90)	1.29 (1.17-1.37)
Negative	0.37 (0.26-0.53)	0.39 (0.27-0.56)	0.25 (0.11-0.58)
AUROC, (95% CI)	0.80 (0.74-0.85)	0.77 (0.71-0.82)	0.65 (0.59-0.70)

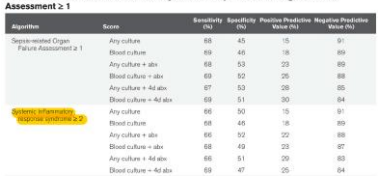
The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

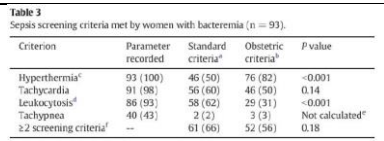
PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if: ☒ Studies inconsistent (*wide variation of treatment effect across*
Tool: Australasian Triage Scale						
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	

Total # of Studies: 1 # of Systematic Reviews: 0 # of RCTs: 0 # of Non-Randomized Studies: 0 # of Diagnostic Studies: 1						
Chamberlain, D. J., et al. (2015). *Emergency Medicine Journal*	To investigate the accuracy and validity of the Australasian Triage Scale (ATS) as a tool to identify and manage in a timely manner the deteriorating patient with severe sepsis.	Prospective cohort study.	**5622 ICU admissions that were screened for sepsis**	The overall sensitivity of th**e ATS to identify severe sepsis was 71% (95% CI 0.676-0.741) with a PPV of 0.586 (95% CI 0.554-0.559)** 	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☐ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	*studies, populations, interventions, or outcomes varied)* ☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)* ☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)* <u>Increase Quality Rating if:</u> ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect

						Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low

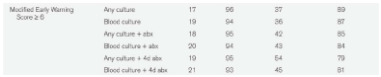
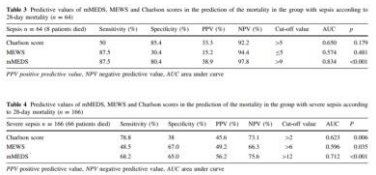
The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if:
Tool: SIRS						☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)* ☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)* ☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	
Total # of Studies: 2 # of Systematic Reviews: 0 # of RCTs: 0 # of Non-Randomized Studies: 0 # of Diagnostic Studies: 2						
Churpek, M. M., et al. (2017). *Critical Care Medicine*	To determine the impact of differing criteria on the accuracy of sepsis screening tools and early warning scores.	Prospective cohort study of all **patients hospitalized outside of the ICU at an academic medical center**.	150,288 patients were hospitalized and 53,849 met at least one infection criteria.	***Accuracy for predicting composite outcome (death+ ICU admissions) across all criteria of infection*** AUC 0.60 (95% CI 0.60-0.61) Sensitivity of ≥2 ranged from 66%-69% TABLE 2. Accuracy of the Different Scores Across Criteria for Suspicion of Infection, With Cut-Offs Matched to the Sensitivity of Quick Sepsis-Related Organ Failure Assessment ≥ 1 	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis
Maguire, P. J., et al. (2016). *International Journal of*	To examine, in the setting of maternal bacteremia, the implications for	**Retrospective cohort study of women with maternal bacteremia**	93 women included	Of 93 women with bacteremia, 61 (66%) would have been diagnosed with sepsis based on standard SIRS compared with 52 (56%) based on customized SIRS (P=0.18).	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner	

Gynaecology & Obstetrics	the diagnosis of maternal sepsis of customizing the systemic inflammatory response syndrome (SIRS) criteria for physiologic changes of pregnancy.			 **Table 3** Sepsis screening criteria met by women with bacteremia (n = 93). Criterion / Parameter recorded / Standard criteria[a] / Obstetric criteria[b] / *P* value Hyperthermia[c] 93 (100) 46 (50) 76 (82) <0.001 Tachycardia 91 (98) 56 (60) 46 (50) 0.14 Leukocytosis[d] 86 (93) 58 (62) 29 (31) <0.001 Tachypnea 40 (43) 2 (2) 3 (3) Not calculated[e] ≥2 screening criteria[f] -- 61 (66) 52 (56) 0.18	☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	*(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)* Increase Quality Rating if: ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if:
Tool: MEWS						☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)* ☒ Studies are indirect *(PICO question is quite different from the*
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	
Total # of Studies: 2 # of Systematic Reviews: 0 # of RCTs: 0 # of Non-Randomized Studies: 0 # of Diagnostic Studies: 2						
Churpek, M. M., et al. (2017). *Critical Care Medicine*	To determine the impact of differing criteria on the accuracy of	Prospective cohort study of all patients hospitalized outside of the ICU at an academic medical center.	150,288 patients were hospitalized and 53,849 met at least one infection criteria.	***Accuracy for predicting composite outcome (death+ ICU admissions) across all criteria of infection***	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study	

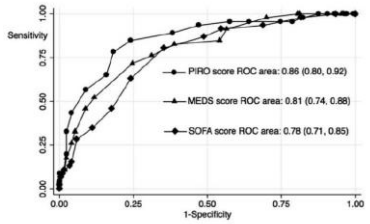
	sepsis screening tools and early warning scores.			AUC 0.65 (95% CI 0.62-0.68) Specificity ranged from 51-57% at the qSOFA cutoff >1  At qSOFA >2 the sensitivity ranged from 17-21% 	☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	*available evidence in regard to population, intervention, comparison, or outcome)* ☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)* <u>Increase Quality Rating if:</u> ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low
Cildir, E., et al. (2013). *Internal & Emergency Medicine*	To evaluate the ability of the modified mortality in emergency department sepsis (MEDS) score, the modified early warning score (MEWS) and the Charlson comorbidity index (CCI) to predict prognosis in patients who are diagnosed in sepsis.	Prospective cohort study of the CCI, MEWS and modified MEDS Score in the prediction of 28-day mortality in **patients presenting to the emergency department who were diagnosed with sepsis.**	230 patients	***Predictive value of 28-day mortality in patients with sepsis***: sensitivity 87.5%, specificity 30.4% AUC 0.574 ***Predictive value of 28-day mortality in patients with severe sepsis***: sensitivity 48.5% specificity 67.0% AUC 0.596 	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

The GRADE criteria were

used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?

Tool: SOFA vs MEDS

Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*
Total # of Studies: 3 # of Systematic Reviews: 0 # of RCTs: 0 # of Non-Randomized Studies: 0 # of Diagnostic Studies: 3					
Macdonald, S. P., et al. (2014). *Academic Emergency Medicine*	To predict mortality in ED patients with features suggesting severe sepsis or septic shock in the ED using MEDS and SOFA scores.	Prospective cohort study **of septic patients in the ED.**	240 patients	AUC for mortality was 0.81 (95% CI = 0.74 to 0.88) for MEDS, and 0.78 (95% CI = 0.71 to 0.87) for SOFA scores. **Pairwise comparisons of the AUC MEDS versus SOFA was not statistically significant (p = 0.37).**  **Figure 2.** ROC curves for PIRO, MEDS, and SOFA scores for 30-day mortality (n = 240). MEDS = Mortality in Emergency Department Sepsis; PIRO = Predisposition Insult Response and Organ failure; ROC = receiver operator characteristic; SOFA = Sequential Organ Failure Assessment.	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis
Wang, J. Y., et al. (2016). American Journal of Emergency Medicine	To investigate the performance of qSOFA in predicting mortality and intensive care	Retrospective cohort study	477 patients clinically **diagnosed with infection in the ED**	The area under the receiver operating characteristic curve (AUC) of **MEDS** (**0.751**) is similar to that of **SOFA (0.729)** and APACHE II (0.732) in predicting 28-day mortality. The AUC of qSOFA, **SOFA**, **MEDS**, and APACHE II in predicting ICU	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study

Lower Quality Rating if:

☐ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)*

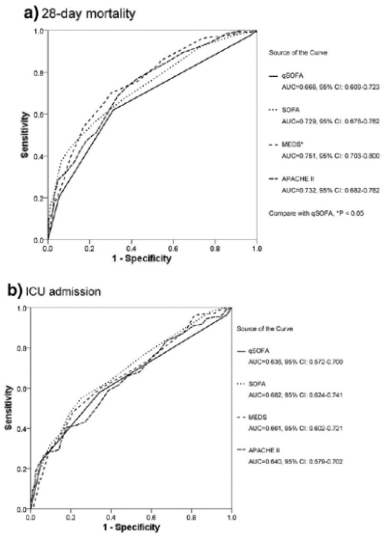
☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)*

☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)*

☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)*

Increase Quality Rating if:

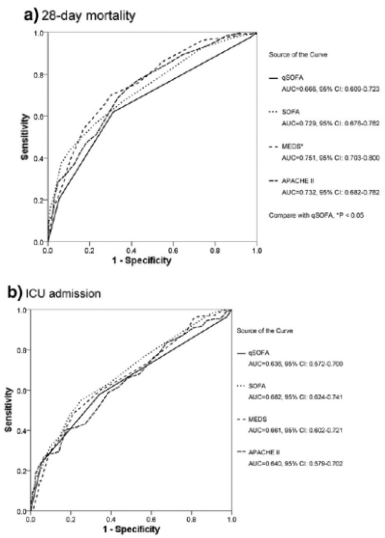
☐ Large Effect

	unit (ICU) admission in patients with clinically diagnosed infection and to compare its performance with that of **MEDS**, APACHE II, and **SOFA**.			admission were 0.636, **0.682, 0.661**, and 0.640, respectively. **There were no significant differences among the score systems.** 	☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☒ Moderate ☐ Low ☐ Very Low
Williams, J. M., et al. (2016). *Critical Care Medicine*	The objectives of this study were to 1) validate a number of severity of illness scores in a large cohort of emergency department patients admitted with presumed	Prospective cohort study of **patients admitted with presumed infection**	8,871 patients	**AUC values for the entire cohort were: MEDS score of 0.92, SOFA score of 0.86.** All scores overestimated mortality, but closest concordance between predicted and observed mortality was seen with Mortality in Emergency Department Sepsis (MEDS) score.	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition	

	infection and 2) compare the performance of scores in patient subgroups with increasing mortality: infection without systemic inflammatory response syndrome, sepsis, severe sepsis, and septic shock.				☐ Failure to include all patients in analysis	

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if:
Tool: APACHE II						☐ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)*
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)*
Total # of Studies: 2 # of Non-Randomized Studies: 2						☐ Studies are imprecise (*When*
Wang, J. Y., et al. (2016). American Journal of Emergency Medicine	To investigate the performance of qSOFA in predicting mortality and intensive care unit (ICU) admission in patients with clinically diagnosed infection and to compare its performance with	Retrospective cohort study	477 patients clinically **diagnosed with infection in the ED**	The area under the receiver operating characteristic curve of qSOFA was lower than that of MEDS (0.666 vs 0.751; P< .05) and similar to that of SOFA (0.729) and **APACHE II (0.732) in predicting 28-day mortality**. The areas under the receiver operating characteristic curve of qSOFA, SOFA, MEDS, and **APACHE II** in predicting ICU admission were 0.636, 0.682, 0.661, and **0.640**, respectively. **There were no significant differences among the score systems.**	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition	

	that of MEDS, **APACHE II**, and SOFA.				☐ Failure to include all patients in analysis	*studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)* <u>Increase Quality Rating if:</u> ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☒ Moderate ☐ Low ☐ Very Low
Williams et al., 2016, *Critical Care Medicine*	To 1) validate a number of severity of illness scores in a large cohort of emergency department patients admitted with presumed infection and 2) compare the performance of scores in patient subgroups with increasing mortality: infection without systemic inflammatory	Prospective observational study	**8871 emergency department patients admitted with presumed infection**	Area under the receiver operating curve values for the entire cohort for **predicting mortality** were: Mortality in Emergency Department Sepsis score of 0.92, **Simplified Acute Physiology Score 11 and Acute Physiology and Chronic Health Evaluation II scores of 0.90,** Sequential Organ Failure Assessment score of 0.86, and Severe Sepsis Score of 0.82. Discrimination decreased in subgroups with greater mortality for each score. All scores overestimated mortality, but closest concordance between predicted and observed mortality was seen with Mortality in Emergency DepartmentSepsis score.	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

	response syndrome, sepsis, severe sepsis, and septic shock					

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if:
Tool: Step-by-Step						
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	
Total # of Studies: 1; # of Non-Randomized Studies: 1						
Gomez, 2016, *Pediatrics*	To validate the Step-by-Step approach and compare it with the Rochester criteria and the Lab-score	Prospective observational study ; compared accuracy of the Step-by-Step approach, the Rochester criteria and the Lab-score in identifying patients at low risk of IBI (isolation of bacterial pathogen in a blood or cerebrospinal fluid)	**2185 nfants </= 90 days with fever without source presenting in 11 European EDs**	Eighty-seven of 2185 infants (4.0%) were diagnosed with an IBI. The prevalence of IBI was significantly higher in infants classified as high risk or intermediate risk according to the Step by Step than in low risk patients. **Sensitivity and negative predictive value for ruling out an IBI were 92.0% and 99.3% for the Step by Step, 81.6% and 98.3% for the Rochester criteria, and 59.8% and 98.1%for the Lab-score.** Seven infants with an IBI were misclassified by the Step by Step, 16 by Rochester criteria, and 35 by the Lab- score	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)* ☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)* ☐ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness*

						of drug, only small, positive studies found) Increase Quality Rating if: ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if: ☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)* ☐ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention,*
Tool: NEWS						
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	
Total # of Studies: 1 # of Non-Randomized Studies: 1						
Keep et al. 2015. *Emerg Med J*	To look at the relationship between the initial national EWS (NEWS) in the emergency department (ED) and the diagnosis of	Retrospective observational study	500 consecutive **non-trauma adult patients presenting to the ED with Manchester Triage System (MTS) category 1- 3**	The area under the curve (AUC) for NEWS to identify patient at risk for SS is 0.89 (95% CI 0.84 to 0.94). A NEWS of 3 or more at ED triage has a sensitivity of 92.6% (95% CI 74.2% to 98.7%) and a specificity of 77% (95% CI 72.8% to 80.6%) to **detect patients at risk for SS** at ED triage.	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test	

	severe sepsis and septic shock (SS)				☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	*comparison, or outcome)* ☒ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)* <u>Increase Quality Rating if:</u> ☐ Large Effect ☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix A.

PICO Question: For hospitalized or ED patients, what tool (SIRS, qSOFA, MEWS, etc) is most effective in identifying sepsis early?						Lower Quality Rating if: ☒ Studies inconsistent (*wide variation of treatment effect across studies, populations, interventions, or outcomes varied)* ☒ Studies are indirect *(PICO question is quite different from the available evidence in regard to population, intervention, comparison, or outcome)* ☒ Studies are imprecise (*When studies include few patients and few events and thus have wide confidence intervals and the results are uncertain)* ☐ Publication Bias *(e.g. pharmaceutical company sponsors study on effectiveness of drug, only small, positive studies found)* Increase Quality Rating if: ☐ Large Effect
Tool: MISSED						
Author/Date	*Purpose of Study*	*Study Design & Methods*	*Sample*	*Outcomes*	*Design Limitations*	
Total # of Studies: 1 # Non-Randomized Studies: 1						
Sivayohama, 2015, European Journal of Emergency Medicine	To validate the simplified MISSED score for predicting all-cause mortality in the ED population admitted with sepsis. The secondary end-point is to validate the risk stratification for ICU admission.	Retrospective observational study	674 **ED patients admitted with a diagnosis of sepsis**	**The area under the curve for predicting mortality for the simplified MISSED score was 0.74** [95% confidence interval (CI) 0.70- 0.77; P< 0.00011. The test characteristics for mortality were as follows: sensitivity 93.9% (95% CI 85-98.3), specificity 37.9% (95% CI 34.1-41.9), positive predictive value 13.9% (95% CI 10.8-17.5), and negative predictive value 98.3% (95% CI 95.7- 99.5). The odds ratio for mortality for a score 2 or more was 5.01 (95% CI 2.93-8.57; P < 0.0001), and that for ICU admission was 3.00 (95% CI 1.70-5.28; P= 0.0001).	Study Limitations = ☐ None ***Diagnostic Studies*** ☐ Patients not enrolled in consecutive or random manner ☐ Case-control study ☒ No independent, blind comparison between index test and reference test ☒ Not all patients received reference test ☒ Reference test not likely to correctly classify target condition ☐ Failure to include all patients in analysis	

						☐ Dose-response gradient ☐ Plausible confounders or other biases increase certainty of effect Quality (certainty) of evidence for studies as a whole: ☐ High ☐ Moderate ☒ Low ☐ Very Low

The GRADE criteria were used to evaluate the quality of evidence presented in research articles reviewed during the development of this guideline. For more detailed information, see Appendix

Guideline Recommendations:

Four guidelines were found that provide recommendations related to the identification of sepsis.

The 2017 **Surviving Sepsis Campaign**: International Guidelines for Management of Sepsis and Septic Shock recommend:

That hospitals and hospital systems have a performance improvement program for sepsis, including sepsis screening for acutely ill, high-risk patients.

Best Practice Statement

The **National Institute for Health and Clinical Excellence** (NICE)'s 2016 guideline on Recognition, Diagnosis and Early Management of Sepsis recommends the following:

Face-to-face Assessment of People with Suspected Sepsis

Assess temperature, heart rate, respiratory rate, blood pressure, level of consciousness and oxygen saturation in young people and adults with suspected sepsis.

Assess temperature, heart rate, respiratory rate, level of consciousness, oxygen saturation and capillary refill time in children under 12 years with suspected sepsis.

Measure blood pressure of children under 5 years if heart rate or capillary refill time is abnormal and facilities to measure blood pressure, including a correctly-sized blood pressure cuff, are available.

Measure blood pressure of children aged 5 to 11 years who might have sepsis if facilities to measure blood pressure, including a correctly-sized cuff, are available.

Examine people with suspected sepsis for mottled or ashen appearance, cyanosis of the skin, lips or tongue, non-blanching rash of the skin, any breach of skin integrity (for example, cuts, burns or skin infections) or other rash indicating potential infection.

Ask the person, parent or carer about frequency of urination in the past 18 hours.

Risk stratification tool for adults, children and young people aged 12 years and over with suspected sepsis (from NICE guideline NG51)

Category	High risk criteria	Moderate to high risk criteria	Low risk criteria
History	Objective evidence of new altered mental state	History from patient, friend or relative of new onset of altered behaviour or mental state History of acute deterioration of functional ability	Normal behaviour

		Impaired immune system (illness or drugs including oral steroids) Trauma, surgery or invasive procedures in the last 6 weeks	
Respiratory	Raised respiratory rate: 25 breaths per minute or more New need for oxygen (40% FiO_2 or more) to maintain saturation more than 92% (or more than 88% in known chronic obstructive pulmonary disease)	Raised respiratory rate: 21–24 breaths per minute	No high risk or moderate to high risk criteria met
Blood pressure	Systolic blood pressure 90 mmHg or less or systolic blood pressure more than 40 mmHg below normal	Systolic blood pressure 91–100 mmHg	No high risk or moderate to high risk criteria met
Circulation and hydration	Raised heart rate: more than 130 beats per minute Not passed urine in previous 18 hours. For catheterised patients, passed less than 0.5 ml/kg of urine per hour	Raised heart rate: 91–130 beats per minute (for pregnant women 100–130 beats per minute) or new onset arrhythmia Not passed urine in the past 12–18 hours For catheterised patients, passed 0.5–1 ml/kg of urine per hour	No high risk or moderate to high risk criteria met
Temperature		Tympanic temperature less than 36°C	
Skin	Mottled or ashen appearance Cyanosis of skin, lips or tongue Non-blanching rash of skin	Signs of potential infection, including redness, swelling or discharge at surgical site or breakdown of wound	No non-blanching rash

Risk stratification tool for children aged 5–11 years with suspected sepsis (from NICE guideline

Category	Age	High risk criteria	Moderate to high risk criteria	Low risk criteria
Behaviour	**Any**	Objective evidence of altered behaviour or mental state Appears ill to a healthcare professional Does not wake or if roused does not stay awake	Not behaving normally Decreased activity Parent or carer concern that the child is behaving differently from usual	Behaving normally

Respiratory	**Any**	Oxygen saturation of less than 90% in air or increased oxygen requirement over baseline	Oxygen saturation of less than 92% in air or increased oxygen requirement over baseline	No high risk or moderate to hig risk criteria met
Aged 5 years		Raised respiratory rate: 29 breaths per minute or more	Raised respiratory rate: 24–28 breaths per minute	
Aged 6–7 years		Raised respiratory rate: 27 breaths per minute or more	Raised respiratory rate: 24–26 breaths per minute	
Aged 8–11 years		Raised respiratory rate: 25 breaths per minute or more	Raised respiratory rate: 22–24 breaths per minute	
Circulation and hydration	**Any**	Heart rate less than 60 beats per minute	Capillary refill time of 3 seconds or more Reduced urine output For catheterised patients, passed less than 1 ml/kg of urine per hour	No high risk or moderate to hig risk criteria met
Aged 5 years		Raised heart rate: 130 beats per minute or more	Raised heart rate: 120–129 beats per minute	
Aged 6–7 years		Raised heart rate: 120 beats per minute or more	Raised heart rate: 110–119 beats per minute	
Aged 8–11 years		Raised heart rate: 115 beats per minute or more	Raised heart rate: 105–114 beats per minute	
Temperature	**Any**		Tympanic temperature less than 36°C	
Skin	**Any**		Mottled or ashen appearance Cyanosis of skin, lips or tongue Non-blanching rash of skin	
Other	**Any**		Leg pain Cold hands	

Risk stratification tool for children aged 5–11 years with suspected sepsis (from NICE guideline

Category	Age	High risk criteria	Moderate to high risk criteria	Low risk criteria
Behaviour	**Any**	No response to social cues Appears ill to a healthcare professional Does not wake, or if roused does not stay awake Weak high-pitched or continuous cry	Not responding normally to social cues No smile Wakes only with prolonged stimulation Decreased activity Parent or carer concern that child is behaving differently from usual	Responds normally to social cues Content or smiles Stays awake or awakens quickly Strong normal cry or not crying

Respiratory	**Any**	Grunting Apnoea Oxygen saturation of less than 90% in air or increased oxygen requirement over baseline	Oxygen saturation of less than 92% in air or increased oxygen requirement over baseline Nasal flaring	No high risk or moderate to high risk criteria met
Under 1 year		Raised respiratory rate: 60 breaths per minute or more	Raised respiratory rate: 50–59 breaths per minute	
1–2 years		Raised respiratory rate: 50 breaths per minute or more	Raised respiratory rate: 40–49 breaths per minute	
3–4 years		Raised respiratory rate: 40 breaths per minute or more	Raised respiratory rate: 35–39 breaths per minute	
Circulation and hydration	**Any**	Bradycardia: heart rate less than 60 beats per minute	Capillary refill time of 3 seconds or more Reduced urine output For catheterised patients, passed less than 1 ml/kg of urine per hour	No high risk or moderate to high risk criteria met
Under 1 year		Rapid heart rate: 160 beats per minute or more	Rapid heart rate: 150–159 beats per minute	
1–2 years		Rapid heart rate: 150 beats per minute or more	Rapid heart rate: 140–149 beats per minute	

Levels of Evidence not provided

The **National Collaborating Centre for Cancer** (NCC-C) on behalf of the National Institute for Health and Clinical Excellence (NICE)' s 2012 guideline on The Prevention and Management of Neutropenic Sepsis recommends the following:

When to Refer Patients in the Community for Suspected Neutropenic Sepsis

Suspect neutropenic sepsis in patients having anticancer treatment who become unwell. Refer patients with suspected neutropenic sepsis immediately for assessment in secondary or tertiary care.

Managing Suspected Neutropenic Sepsis in Secondary and Tertiary Care

Emergency Treatment and Assessment

Treat suspected neutropenic sepsis as an acute medical emergency and offer empiric antibiotic therapy immediately.

Include in the initial clinical assessment of patients with suspected neutropenic sepsis:

- History and examination
- Full blood count, kidney and liver function tests (including albumin), C-reactive protein, lactate and blood culture

Levels of Evidence not provided

The **Royal College of Obstetricians and Gynaecologists** 2012 guideline for Bacterial Sepsis in Pregnancy recommends the following:

All healthcare professionals should be aware of the symptoms and signs of maternal sepsis and critical illness and of the rapid, potentially lethal course of severe sepsis and septic shock. Suspicion of significant sepsis should trigger an urgent referral to secondary care.

Clinical signs suggestive of sepsis include one or more of the following: pyrexia, hypothermia, tachycardia, tachypnoea, hypoxia, hypotension, oliguria, impaired consciousness and failure to respond to treatment. These signs, including pyrexia, may not always be present and are not necessarily related to the severity of sepsis.

Blood cultures are the key investigation and should be obtained prior to antibiotic administration; however, antibiotic treatment should be started without waiting for microbiology results.

Serum lactate should be measured within six hours of the suspicion of severe sepsis in order to guide management. Serum lactate ≥4 mmol/l is indicative of tissue hypoperfusion.
Any relevant imaging studies should be performed promptly in an attempt to confirm the source of infection.
Evidence Level 4

Guideline Ratings

Guideline Issuer and Date	NICE 2016	Royal College 2012	Surviving Sepsis 2016	NCC-C/NICE 2012
1. Transparency	A	A	A	A
2. Conflict of interest	NR	NR	A	NR
3. Development group	A	A	A	A
4. Systematic Review	A	A	A	A
5. Supporting evidence	B	A	A	B
6. Recommendations	A	A	A	A
7. External Review	A	A	A	A
8. Currency and updates	B	C	A	C

See appendix B for full description of the Trustworthy Guideline grading system.

REFERENCES:

1. Makam, A. N., et al. (2015). "Diagnostic accuracy and effectiveness of automated electronic sepsis alert systems: A systematic review." Journal of Hospital Medicine (Online) 10(6): 396-402.
2. Churpek, M. M., et al. (2017). "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." Critical Care Medicine 45(11): 1805-1812.
3. Chen, Y. X., et al. (2016). "Use of CRB-65 and quick Sepsis-related Organ Failure Assessment to predict site of care and mortality in pneumonia patients in the emergency department: a retrospective study." Critical Care (London, England) 20(1): 167.
4. Wang, J. Y., et al. (2016). "Predictive performance of quick Sepsis-related Organ Failure Assessment for mortality and ICU admission in patients with infection at the ED." American Journal of Emergency Medicine 34(9): 1788-1793.
5. Kim, M., et al. (2017). "Predictive performance of the quick Sequential Organ Failure Assessment score as a screening tool for sepsis, mortality, and intensive care unit admission in patients with febrile neutropenia." Supportive Care in Cancer 25(5): 1557-1562.
6. Moskowitz, A., et al. (2017). "Quick Sequential Organ Failure Assessment and Systemic Inflammatory Response Syndrome Criteria as Predictors of Critical Care Intervention Among Patients With Suspected Infection." Critical Care Medicine 45(11): 1813-1819.
7. Askim, A., et al. (2017). "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." Scandinavian Journal of Trauma, Resuscitation & Emergency Medicine 25(1): 56.
8. Williams, J. M., et al. (2017). "Systemic Inflammatory Response Syndrome, Quick Sequential Organ Function Assessment, and Organ Dysfunction: Insights From a Prospective Database of ED Patients With Infection." Chest 151(3): 586-596.
9. Freund, Y., et al. (2017). "Prognostic Accuracy of Sepsis-3 Criteria for In-Hospital Mortality Among Patients With Suspected Infection Presenting to the Emergency Department." JAMA 317(3): 301-308.
10. Chamberlain, D. J., et al. (2015). "Identification of the severe sepsis patient at triage: a prospective analysis of the Australasian Triage Scale." Emergency Medicine Journal 32(9): 690-697.
11. Maguire, P. J., et al. (2016). "Evaluation of the systemic inflammatory response syndrome criteria for the diagnosis of sepsis due to maternal bacteremia." International Journal of Gynaecology & Obstetrics 133(1): 116-119.
12. Cildir, E., et al. (2013). "Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community acquired sepsis in the emergency department." Internal & Emergency Medicine 8(3): 255-260.
13. Macdonald, S. P., et al. (2014). "Comparison of PIRO, SOFA, and MEDS scores for predicting mortality in emergency department patients with severe sepsis and septic shock." Academic Emergency Medicine 21(11): 1257-1263.
14. Williams, J. M., et al. (2016). "Severity Scores in Emergency Department Patients With Presumed Infection: A Prospective Validation Study." Critical Care Medicine 44(3): 539-547.

15. Gomez, B., et al. (2016). "Validation of the "Step-by-Step" Approach in the Management of Young Febrile Infants." Pediatrics 138(2).
16. Keep, J. W., et al. (2016). "National early warning score at Emergency Department triage may allow earlier identification of patients with severe sepsis and septic shock: a retrospective observational study." Emergency Medicine Journal 33(1): 37-41.
17. Sivayoham, N., et al. (2015). "Internal emergency department validation of the simplified MISSED score." European Journal of Emergency Medicine 22(5): 321-326.

Appendix A. GRADE criteria for rating a body of evidence on an intervention

Developed by the GRADE Working Group

Grades and interpretations:

High: Further research is very unlikely to change our confidence in the estimate of effect. Moderate: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. Low: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Very low: Any estimate of effect is very uncertain.

Type of evidence and starting level

Randomized trial–high Observational study–low Any other evidence–very low

Criteria for increasing or decreasing level

Reductions Study quality has serious (–1) or very serious (–2) problems Important inconsistency in evidence (–1) Directness is somewhat (–1) or seriously (–2) uncertain Sparse or imprecise data (–1) Reporting bias highly probable (–1) **Increases** Evidence of association† strong (+1) or very strong (+2) †Strong association defined as significant relative risk (factor of 2) based on consistent evidence from two or more studies with no plausible confounders Very strong association defined as significant relative risk (factor of 5) based on direct evidence with no threats to validity.

Appendix B. Trustworthy Guideline rating scale

The University of Pennsylvania's Center for Evidence-Based Practice Trustworthy Guideline rating scale is based on the Institute of Medicine's "Standards for Developing Trustworthy Clinical Practice Guidelines" (IOM), as well as a review of the AGREE Enterprise and Guidelines International Network domains.

The purpose of this scale is to focus on the weaknesses of a guideline that may reduce the trust a clinical user can have in the guideline, and distinguish weaknesses in documentation (e.g. guide-line does not have a documented updating process) from weaknesses in the guidance itself (e.g. recommendations are outdated). Current quality scales like AGREE emphasize documentation. They are important checklists for developers of new guidelines, but are less useful for grading existing guidelines. These scales also are harder for clinicians and other persons who are not methodology experts to apply, and their length discourages their use outside formal technology assessment reports. This new scale is brief, balanced, and easy and consistent to apply.

We do not attempt to convert the results of this assessment into a numeric score. Instead we present a table listing the guidelines and how they are rated on each standard. This facilitates qualitative understanding by the reader, who can see for what areas the guideline base as a whole is weak or strong as well as which guidelines are weaker or stronger.

1. Transparency

A	Guideline development methods are fully disclosed.
B	Guideline development methods are partially disclosed.
C	Guideline development methods are not disclosed.

The grader must refer to any cited methods supplements or other supporting material when evaluating the guideline. Methods should include:

Who wrote the initial draft
How the committee voted on or otherwise approved recommendations

Evidence review, external review and methods used for updating are not addressed in this standard.

2. Conflict of interest

A	Funding of the guideline project is disclosed, disclosures are made for each individual panelist, and financial or other conflicts do not apply to key authors of the guideline or to more than 1 in 10 panel members).
B	Guideline states that there were no conflicts (or fewer than 1 in 10 panel members), but does not disclose funding source.

C	Lead author, senior author, or guideline panel members (at least 1 in 10) have conflict of interest, or guideline project was funded by industry sponsor with no assurance of independence.
NR	Guideline does not report on potential conflict of interests.

For purposes of this checklist, conflicts of interest include employment by, consulting for, or holding stock in companies doing business in fields affected by the guideline, as well as related financial conflicts. This definition should not be considered exclusive. As much as anything, this is a surrogate marker for thorough reporting, since it may be assumed that guideline projects are funded by the sponsoring organization and many authors think it unnecessary to report a non-conflict.

3. Guideline development group

A	Guideline development group includes 1) methodological experts and clinicians and 2) representatives of multiple specialties.
B	Guideline development group includes one of the above, but not both.
C	Guideline developers all from one specialty or organization, and no methodologists.
NR	Affiliations of guideline developers not reported

The purpose of this standard is to ensure that supporters of competing procedures, or clinicians with no vested interest in utilization of one procedure or another, are involved in development of the guideline. Both AGREE II and IOM call for patient or public involvement: very few guideline panels have done so to date, so this is not necessary for guidelines to be rated A. Involvement of methodologists or HTA specialists in the systematic review is sufficient involvement in the guideline development group for our purposes. In the absence of any description of the guideline group, assume the named authors are the guideline group.

4. Systematic review

A	Guideline includes a systematic review of the evidence or links to a current review.
B	Guideline is based on a review which may or may not meet systematic review criteria.
C	Guideline is not based on a review of the evidence.

In order to qualify as a systematic review, the review must do all of the following:

Describe itself as systematic or report search strategies using multiple databases
Define the scope of the review (including key questions and the applicable population)
Either include quantitative or qualitative synthesis of the data or explain why it is not indicated

Note: this element does not address the quality of the systematic review: simply whether or not it exists. Concerns about quality or bias of the review will be discussed in text, where the analyst will explain whether the weaknesses of the review weaken the validity or reliability of the guideline.

Note: a guideline may be rated B on this domain even if the review on which it is based is not available to us. This potential weakness of the guideline should be discussed in text of the report.

5. Grading the supporting evidence

A	Specific supporting evidence (or lack thereof) for each recommendation is cited and graded
B	Specific supporting evidence (or lack thereof) for each recommendation is cited but the recommendation is not graded.
C	Recommendations are not supported by specific evidence.

To score a B on this domain there should be specific citations to evidence tables or individual references for each relevant recommendation in the guideline, or an indication that no evidence was available. Any standardized grading system is acceptable for purposes of this rating. If a guideline reports that there is no evidence available despite a thorough literature search, it may be scored B on this domain, or even A if evidence for other recommendations is cited and graded.

6. Recommendations

A	Considerations for each recommendation are documented (i.e. benefits and harms of a particular action, and/or strength of the evidence); and recommendations are presented in an actionable form.
B	Either one or the other of the above criteria is met.
C	Neither of the above criteria are met

In order to be actionable, the guideline should specify the specific population to which the guideline applies, the specific intervention in question, and the circumstances under which it should be carried out (or not carried out). The language used in the recommendations should also be consistent with the strength of the recommendation (e.g. directive and active language like “should” or “should not” for strong recommendations, and passive language like “consider” for weak recommendations). A figure or algorithm is considered actionable as long as it is complete enough to incorporate all the applicable patients and interventions. Please see the forthcoming NICE manual (24) for a good discussion of actionability in guidelines.

7. External review

A	Guideline was made available to external groups for review.

B	Guideline was reviewed by members of the sponsoring body only.
C	Guideline was not externally reviewed.
NR	No external review process is described.

8. Updating and currency of guideline

A	Guideline is current and an expiration date or update process is specified.
B	Guideline is current but no expiration date or update process is specified.
C	Guideline is outdated.

A guideline is considered current if it is within the developers' stated validity period, or if no period or expiration data is stated, the guideline was published in the past three years (NOTE: the specific period may be changed at the analyst's discretion, based on whether the technology is mature and whether there is a significant amount of recent evidence). A guideline must address new evidence when it is updated. A guideline which is simply re-endorsed by the panel without searching for new evidence must be considered outdated