Sepsis Screening Tools in Acute Care: A Comparative Narrative Review of Diagnostic Accuracy, Clinical Performance, and Implementation Effectiveness (2015–2025)

G Prasad*
Department of Trauma & Emergency, All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

Harideep Yellamilli
Department of Trauma & Emergency, All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

Anjani Priya V
Department of Trauma & Emergency, All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

Muni Srikanth I
Department of Orthopaedics, All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

Rohini Dattatri
Department of Trauma & Emergency, All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

Cheranjeevi J
Department of Dentistry, All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

TVN Sriranjitha
Department of Transfusion Medicine, Siddhartha Medical College Vijayawada, Andhra Pradesh 520008, India

Chaitanya Kumar IS
Department of Transfusion Medine & Hemotherapy All India Institute of Medical Sciences, Mangalagiri, Andhra Pradesh 522503, India

Editorial history

Received: 15.01.2026
Accepted: 18.01.2026
Published Online: 10.02.2026

Cite this article

Prasad, G., Yellamilli, H., Cheranjeevi J., et al. (2025). Sepsis Screening Tools in Acute Care: A Comparative Narrative Review of Diagnostic Accuracy, Clinical Performance, and Implementation Effectiveness (2015–2025). *International Journal of Elite Research and Innovation, 1*(1), 01-17.

Abstract

Purpose: Sepsis remains a global health emergency with mortality exceeding 20%, yet early recognition and screening remain inconsistent. This narrative review synthesizes evidence on sepsis screening tools and early warning scores for adult patients in acute-care environments.

Scope of literature: A comprehensive search of PubMed, Embase, Scopus, Web of Science, and Cochrane Library (January 2015–December 2025) identified 26+ primary studies, systematic reviews, meta-analyses, and clinical guidelines. Key sources include 62,338-patient meta-analyses, Sepsis-3 consensus1, Surviving Sepsis Campaign 2021 guidelines, and recent ED implementation trials.

Key themes: (1) No single screening tool offers optimal sensitivity and specificity; (2) qSOFA excels as a mortality predictor but lacks screening sensitivity; (3) NEWS and MEWS provide balanced early warning; (4) SIRS remains too nonspecific for sepsis; (5) combining tools with lactate or capillary refill improves diagnostic accuracy; (6) electronic alert systems reduce mortality and improve bundle adherence; (7) resource limitations and setting heterogeneity drive performance variation.

Clinical relevance: Current evidence supports multimodal screening strategies combining bedside vital-sign scores (NEWS/MEWS), organ-dysfunction assessment (qSOFA/SOFA), biochemical markers (lactate), and structured electronic or clinical decision-support pathways embedded in quality-improvement programs. No single score should be used in isolation. Clinicians must balance diagnostic sensitivity with specificity based on local case-mix, infrastructure, and antimicrobial stewardship concerns.

MeSH-Aligned Keywords

- Sepsis / diagnosis
- Sepsis / mortality
- Critical Illness / diagnosis
- Shock, Septic / diagnosis
- Early Diagnosis
- Severity of Illness Index
- Clinical Decision Rules
- Sensitivity and Specificity
- Receiver Operating Characteristic (ROC) Curve
- Emergency Service, Hospital
- Intensive Care Units
- Quality Improvement
- Electronic Health Records
- Machine Learning
- Alert Systems, Clinical
- Evidence-Based Medicine
- Systematic Reviews as Topic

INTRODUCTION

Problem Definition: The Sepsis Recognition Gap

Sepsis—defined as life-threatening organ dysfunction caused by dysregulated host response to infection—affects approximately 20 million patients annually worldwide, resulting in over 5 million deaths.1,2 Despite advances in antimicrobial therapy, critical care, and supportive measures, sepsis remains among the leading causes of hospital mortality, accounting for one in three to six deaths in hospitalized patients2. In high-income countries, sepsis

mortality remains 15–25%; in low- and middle-income countries (LMICs), it exceeds 50%2. The burden extends beyond acute illness: an estimated 38 million sepsis survivors annually face profound long-term sequelae including physical disability, cognitive impairment, post-traumatic stress, and unemployment3.

The central challenge lies not in treatment—modern sepsis bundles (early antibiotics, fluid resuscitation, vasopressors, lactate monitoring) demonstrably improve outcomes4—but inearly recognition. Delays of even 1–3 hours in sepsis detection and bundle initiation are associated with exponentially increased mortality risk4,5. Yet across emergency departments (EDs), acute wards, and critical-care units, clinicians face a recognition problem: sepsis presents with nonspecific symptoms (fever, malaise, tachycardia, confusion) that overlap substantially with noninfectious inflammation, acute illness, and comorbid conditions. No single clinical bedside sign reliably distinguishes sepsis from mimics.

Why a Review Is Needed Now

The landscape of sepsis screening has evolved dramatically since the 2016 Sepsis-3 consensus1, which downgraded SIRS criteria and proposed quick Sequential Organ Failure Assessment (qSOFA) as a bedside risk-stratification tool. Subsequent validation studies, guideline updates, and clinical trials have revealed critical limitations of qSOFA—particularly its unacceptably low sensitivity for early sepsis detection6,7,8. The 2021 Surviving Sepsis Campaign Guidelines nowstrongly recommend againstqSOFA as a single screening tool, instead endorsing broader early warning scores (NEWS, MEWS), lactate measurement, and structured electronic or manual alert systems.9

Concurrently, hospitals are increasingly deploying electronic decision-support systems and machine-learning-based sepsis alerts, yet their real-world effectiveness, unintended consequences (alert fatigue, overtreatment), and equity implications remain contested. For healthcare systems in resource-constrained settings, the question of which screening approach optimizes sensitivity and feasibility without requiring point-of-care testing or sophisticated EHR infrastructure is urgent and understudied.

A new narrative review is timely because: (1) meta-analyses and head-to-head ED studies have clarified comparative performance of major screening tools within the last 3–5 years; (2) implementation science literature now quantifies the effect of electronic alerts and quality-improvement bundles on mortality and bundle adherence; (3) heterogeneity in patient populations, sepsis definitions, and outcomes across studies has led to apparent contradictions that merit synthesis and reconciliation; and (4) clinical practice variability persists despite guideline recommendations, suggesting a need for evidence-based clarity on context-appropriate screening strategies.

Current Knowledge and Gaps

What is well established:

- Sepsis-3 definition (infection + SOFA ≥2) improves diagnostic specificity and prognostic stratification compared to SIRS-based definitions.1
- qSOFA (range 0–3) is highly specific for mortality in suspected infection but poorly sensitive for early sepsis detection6,7.
- SIRS has high sensitivity but very low specificity, making it a poor mortality predictor but reasonable "rule-out" tool6,10.
- NEWS and MEWS demonstrate balanced sensitivity and specificity across ED and ward populations, outperforming SIRS and qSOFA in some studies.7,11
- Combining physiological scores with biochemical markers (lactate) and perfusion assessments (capillary refill time) can improve diagnostic discrimination12,13.
- Electronic sepsis alert systems are associated with reduced mortality, shorter hospital stay, and improved guideline adherence14.

Persistent knowledge gaps:

- Why and when does SIRS/qSOFA performance diverge across patient populations? (Answered partly by recent meta-regression; literature heterogeneity remains.)
- Which combination of tools (e.g., NEWS + lactate, or local triage score + qSOFA) optimizes outcomes in specific settings (ED, general ward, ICU, LMIC)?
- What is the true magnitude of alert fatigue and overtreatment harm from widespread electronic sepsis alerts?
- How can screening tools be adapted for low-resource settings where laboratory infrastructure is limited?
- Do dynamic (serial) scores outperform single time-point assessments, and if so, at what cost in workflow and usability?

METHODS

Study Design and Scope

This is narrative review synthesizing evidence on sepsis screening and early warning tools in adult acute-care settings. Narrative reviews permit a structured but comprehensive synthesis of heterogeneous literature (observational cohorts, randomized trials, systematic reviews, implementation studies).This narrative review was conducted and reported in accordance with accepted methodological guidance for narrative literature reviews, emphasizing transparency, reproducibility of search strategy, and critical synthesis of heterogeneous evidence15 to provide clinically nuanced guidance that systematic reviews may not accommodate. The review does not conform to PRISMA-P; however, the search strategy is transparent and reproducible to ensure methodological rigor.

Databases Searched

Comprehensive searches were conducted in:

- PubMed/MEDLINE (via PubMed Central and indexed journal archives)
- Embase (Elsevier)
- Scopus (Elsevier)
- Web of Science Core Collection (Clarivate)
- Cochrane Library (Cochrane Reviews and Central Register of Controlled Trials)

Time Frame and Language Restrictions

- Temporal scope: January 1, 2015, through December 31, 2025.
- Rationale for 2015 start: Captures post-Sepsis-3 consensus (February 2016) and the maturing literature on NEWS, qSOFA, and electronic alert systems.
- Language: English-language publications only.

Search Terms and Strategy

Core search strategy combined Boolean operators and MeSH headings:

- Sepsis AND (screening OR screening tool OR screening score OR early warning score OR triage)
- qSOFA OR "quick SOFA" OR quick Sequential Organ Failure Assessment

- NEWS OR "National Early Warning Score" OR MEWS OR "Modified Early Warning Score"
- SIRS OR "Systemic Inflammatory Response Syndrome
- sepsis AND (alert OR alert system OR electronic alert OR machine learning)
- sepsis AND (lactate OR "capillary refill" OR perfusion OR organ dysfunction
- sepsis AND (emergency department OR ED OR ward OR hospital OR intensive care)
- sepsis AND (diagnostic accuracy OR sensitivity OR specificity OR ROC)

Searches were conducted iteratively with results cross-checked and refined through November 2025.

Inclusion Criteria

Articles were included if they:

1. **Studied adult human populations:** (age ≥18 years) with suspected infection, sepsis, or septic shock.
2. **Evaluated performance of one or more screening/early warning tools:** qSOFA, SIRS, SOFA, mSOFA, NEWS, NEWS2, MEWS, shock index, modified shock index, institutional triage scores, or machine-learning-based sepsis prediction models.
3. **Reported diagnostic test accuracy metrics:** sensitivity, specificity, area under the receiver operating characteristic (AUC-ROC), positive/negative predictive values, likelihood ratios, and OR.
4. **Included outcome measures:** mortality (28-day, 30-day, in-hospital), ICU admission, length of stay, and sepsis bundle adherence.
5. **Represented primary research:** prospective or retrospective observational cohorts, quasi-experimental quality-improvement studies, cluster-randomized trials, diagnostic accuracy studies, or systematic reviews/meta-analyses.
6. **Studied acute-care settings:** emergency departments, acute medical/surgical wards, intensive care units, or prehospital/out-of-hospital care.

Exclusion Criteria

Articles were excluded if they:

- Focused solely on pediatric (<18 years) or neonatal populations.
- Reported purely pathophysiologic, biomarker discovery, or mechanistic studies without evaluating a defined screening tool.
- Consisted of case reports, narrative commentaries, editorials, or conference abstracts without full-text peer-reviewed publication.
- Used animal models or in vitro studies.
- Focused on disease other than sepsis as primary outcome (e.g., trauma, acute coronary syndrome, unless sepsis was a pre-specified analysis).

Article Types Included

- **Original research:** Prospective cohort studies, retrospective cohort studies, cross-sectional surveys.
- **Intervention studies:** Cluster-randomized trials, quasi-experimental quality-improvement studies, and before-and-after designs evaluating electronic alert or bundle implementation.

- **Diagnostic accuracy studies:** ROC analyses, diagnostic test comparison studies.
- **Systematic reviews and meta-analyses:** Reviews comparing two or more screening tools or synthesizing the effectiveness of alert systems.
- **Clinical practice guidelines:** Sepsis-3 consensus definitions, Surviving Sepsis Campaign guidelines, American, European, and international professional society statements.

Approach to Bias and Bias Awareness

This review explicitly acknowledges and discusses biases that may distort evidence interpretation:

1. **Publication bias:** Positive or null findings on screening tool performance may be preferentially published (or unpublished), leading to overestimation of tool accuracy. Conversely, highly negative findings on institutional tools may remain unpublished. This review cites meta-analytic funnel plot assessments and negative studies where available.
2. **Spectrum bias:** Validation cohorts often include patient populations (e.g., ICU-only, high-acuity) that differ from the broader population for which the tool is intended (e.g., all ED presentations). Sensitivity and specificity estimates are skewed when applied across settings.
3. **Verification bias:** Sepsis status is often "verified" by gold standard definitions (Sepsis-2, Sepsis-3, physician judgment) that may differ across studies, limiting comparability.
4. **Study design bias:** Retrospective studies may be subject to incomplete data abstraction, selection bias (e.g., the inclusion of only ICU-admitted patients), and an inability to capture dynamic clinical assessments. Prospective studies are fewer and may have different enrollment practices.
5. **Reference standard heterogeneity:** Different definitions of sepsis (Sepsis-2 vs Sepsis-3, qSOFA ≥2 vs SOFA ≥2 increase, physician-adjudicated vs administrative coding) yield different apparent accuracies for the same screening tool.
6. **Threshold bias:** Cut-off values for qSOFA (≥2 vs ≥1), NEWS (≥5 vs ≥7 vs ≥8), and MEWS vary across studies, affecting sensitivity-specificity trade-offs.
7. **Single-center bias:** Most studies are single-center or limited to one health system, reducing external generalizability.
8. **EHR-dependent bias:** Electronic alert studies may not be reproducible across different EHR platforms or settings with different data availability.

This narrative review addresses these biases by (a) reporting heterogeneity estimates (I^2 values); (b) stratifying analyses by setting, patient type, and outcome definition where possible; (c) highlighting limitations of individual studies; (d) prioritizing meta-analytic syntheses and multicenter validation cohorts; and (e) discussing generalizability explicitly.

RESULTS

Thematic Synthesis

1. Evolution from SIRS to Sepsis-3 and Current Screening Paradigm

SIRS and Early Sepsis Definitions

The Systemic Inflammatory Response Syndrome (SIRS) criteria—introduced by the American College of Chest Physicians and American Society of Critical Care Medicine in 1991—defined sepsis as infection plus ≥2 of: fever/hypothermia, tachycardia (>90 bpm), tachypnea (>20 breaths/min), or abnormal WBC count (>12,000 or <4,000 cells/μL). SIRS was intentionally sensitive (designed to capture patients at risk) but poorly specific for sepsis13, as inflammation from trauma, surgery, autoimmunity, and other noninfectious conditions readily met SIRS criteria.

Subsequent meta-analyses and validation studies demonstrated SIRS's poor prognostic discrimination7: while SIRS identified many infected patients, it failed to distinguish infection-induced sepsis from noninfectious illness and was

a weak predictor of mortality. Yet SIRS remained embedded in clinical practice and, until recently, in guidelines.

Sepsis-3 Consensus and Introduction of qSOFA (2016)

In 2016, a multidisciplinary international panel published the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3), which shifted paradigm emphasis from inflammation (SIRS) toorgan dysfunctionas the defining feature of sepsis1. The Sepsis-3 definition—infection plus acute increase in SOFA score ≥2—reflects the pathophysiology of dysregulated host response manifesting as multi-organ compromise. SOFA (Sequential Organ Failure Assessment) evaluates six organ systems (respiratory, cardiovascular, neurological, hepatic, hematological, renal), each scored 0–4, for a maximum of 24.

Concurrently, Sepsis-3 introduced qSOFA (quick Sequential Organ Failure Assessment), a simplified three-variable bedside score: altered mentation (Glasgow Coma Scale <15; 1 point), elevated respiratory rate (≥22 breaths/min; 1 point), and/or low systolic blood pressure (≤100 mmHg; 1 point). qSOFA ≥2 was presented as a rapid risk-stratification tool for patients with suspected infection outside the ICU, identifying those at high risk of poor outcomes (mortality, prolonged ICU stay).

Key distinction: Sepsis-3 positioned qSOFA as aprognostic tool8 (predicting who will die, not who has sepsis) and recommended a SOFA score≥2 for sepsis diagnosis. However, confusion arose in clinical practice: many interpreted qSOFA as a screening tool for sepsis diagnosis.

Current Meta-Analytic Evidence: SIRS, qSOFA, and NEWS

A 2022 meta-analysis of 26 studies including 62,338 patients with suspected sepsis compared the three tools forpredicting mortality:

qSOFA	46% (95% CI 39–53)	82% (95% CI 76–86)	3.79	39%
SIRS	82% (95% CI 78–85)	24% (95% CI 19–29)	~1.0	21%
NEWS	73% (95% CI 63–81)	52% (95% CI 39–65)	2.96	28%

Key findings:

- **qSOFA** had the highest overall prognostic accuracy (highest diagnostic odds ratio) due to superior specificity, but its low sensitivity (46%) means nearly half of patients who die are not flagged by qSOFA ≥27.
- **SIRS** had high sensitivity (82%) but catastrophically low specificity (24%), making it unable to discriminate sepsis from systemic inflammation. As a sole predictor, SIRS is not recommended.
- **NEWS** offered a middle ground: intermediate sensitivity (73%) and intermediate specificity (52%), avoiding SIRS's low specificity and qSOFA's low sensitivity.

This meta-analysis concluded that qSOFA is superior for mortality prediction but inferior for early screening, and that no single tool provides both high sensitivity and high specificity for sepsis in all populations.

2. NEWS, MEWS, and Early Warning Scores in Emergency Departments

Introduction and Clinical Context

The National Early Warning Score (NEWS) was launched by the UK Royal College of Physicians in 2012 as a generic early-warning system to identify hospitalized patients at risk of deterioration (regardless of underlying diagnosis). NEWS comprises seven physiological parameters:

- Respiratory rate (0–3 points)
- Oxygen saturation SpO_2 (0–3 points)
- Oxygen supplementation required? (Yes/No)
- Temperature (0–3 points)
- Systolic blood pressure (0–3 points)
- Heart rate (0–3 points)

- Level of consciousness (AVPU: Alert/Verbal/Pain/Unresponsive; 0 or 3 points)

The maximum score is 20. A NEWS ≥7 traditionally triggers a high-risk response; NEWS ≥5 is considered moderate risk. NEWS2 (launched 2017) refined SpO_2 scoring to accommodate hypoxemia in chronic lung disease.

The Modified Early Warning Score (MEWS), developed earlier, uses a subset of vital signs and level of consciousness (maximum 14 points) with a threshold of ≥5 triggering escalation.

Recent Comparative ED Studies

Two large recent cohorts directly compared qSOFA, SIRS, NEWS, MEWS, and supplementary measures in ED patients with suspected infection.

Study 1: Oi et al. (2025)12, Japan; n=75 ED patients with suspected infection:

qSOFA	0.833	66.7%	100%	≥2
NEWS	0.846	69%	93%	>7.5
MEWS	0.846	75%	93%	>4.5
SIRS	0.610	81.2%	40.7%	≥2
qSOFA + lactate	0.844	87.5%	81.4%	Combined

This small but rigorous Japanese study found that NEWS and MEWS had statistically equivalent and superior AUCs compared to qSOFA and SIRS alone. Critically, combining qSOFA with lactate measurement elevated both sensitivity and specificity above 80%, addressing qSOFA's screening weakness**.**

Study 2: Wang et al. (2025), CETAT database16[China] + MIMIC-IV [USA]; n=12,799 ED admissions:

Raw screening tool performance for predicting sepsis (infection + organ dysfunction per Sepsis-3):

qSOFA	0.64	0.64
SIRS	0.71	0.55
NEWS	0.74	0.65

This large multinational cohort showed NEWS outperforming qSOFA and SIRS for sepsis detection. Notably, when enhanced through logistic regression incorporating additional clinical variables (age, vitals, lab values), an optimized model achieved AUC 0.83 for sepsis prediction and AUC 0.75 for high-risk sepsis—significantly better than any single score. This underscores the value of combining multiple parameters.

Key Synthesis: Scores as Screening vs. Prognostication

For screening (identifying patients with suspected sepsis who need urgent evaluation and bundle initiation): NEWS/MEWS preferred due to higher sensitivity; qSOFA alone is insufficient.

For prognostication (predicting mortality among confirmed sepsis patients), qSOFA is more specific but misses early cases.

Optimal hybrid approach: Use NEWS/MEWS for initial screening (high sensitivity), follow with qSOFA and SOFA to confirm diagnosis and assess prognosis, and measure lactate for risk stratification.

3. Combining Screening Tools with Biochemical Markers and Perfusion Assessment

Recent evidence increasingly supports multimodal screening: pairing physiological scores with biochemical markers (lactate) and/or objective perfusion measures (quantitative capillary refill time).

Lactate as a Sepsis Marker

Serum lactate (>2 mmol/L, or >18 mg/dL) reflects tissue hypoperfusion and anaerobic metabolism18, hallmarks of sepsis and septic shock. The 2021 Surviving Sepsis Campaign guidelines strongly recommend lactate measurement

in all patients with suspected sepsis and have integrated lactate into the 1-hour bundle. Elevated lactate predicts mortality independent of blood pressure or clinical shock status.

Combination Studies

Oi et al. (2025) demonstrated that combining qSOFA with lactate yielded:

- **qSOFA + lactate: AUC 0.844, sensitivity 87.5%, specificity 81.4%**

This combination achieved >80% sensitivity AND specificity—a marked improvement over qSOFA alone (66.7% sensitivity).

Similarly, qSOFA + quantitative capillary refill time (Q-CRT) yielded AUC 0.821 with sensitivity 83.3% and specificity 81.4%, suggesting that objective measures of peripheral perfusion can substitute for or augment lactate when laboratories are unavailable.

Clinical Implication

For resource-limited settings where blood gas analysis is delayed or unavailable, bedside assessment combining qSOFA with visual assessment of perfusion markers (skin mottling, prolonged capillary refill time >2 seconds, cool extremities) can approximate the utility of lactate-based tools.

4. Institution-Specific Triage Tools and Electronic Alert Systems

Bespoke ED Sepsis Protocols

Recognizing that no single generic score fits all settings, many institutions have developed local sepsis triage algorithms that embed screening tools into ED workflows. For example:

- **Shock index** (HR/SBP; cutoff >0.9) triggers sepsis evaluation and early vasopressor readiness. Shock index and modified shock index have been shown to correlate with illness severity and mortality in septic patients and may improve early risk stratification when incorporated into emergency department triage algorithms17
- **Modified shock index** (HR/MAP; cutoff ≥1.0 or ≥1.3) incorporates diastolic blood pressure, improving specificity.
- **Composite ED triage scores** combining vital signs (RR, HR, BP, O_2 sat, temperature) with clinical judgment flags patients for rapid sepsis workup.

Institution-specific tools, when locally validated, often outperform generic qSOFA for CMS SEP-1 bundle compliance (Centers for Medicare & Medicaid Services sepsis reporting requirement), because they are calibrated to local patient populations and available data systems.

Electronic and Machine-Learning-Based Alert Systems

A 2024 systematic review and meta-analysis 11 of 22 studies (19,580 ED patients) examined the real-world effectiveness of electronic and manual sepsis alert systems:

Overall findings:

- **Sepsis alert systems (any type):** RR 0.83 (95% CI 0.72–0.96) for mortality (17% risk reduction).
- **Electronic alerts (EHR-integrated, rule-based, or ML-based):** RR 0.78 (95% CI 0.67–0.92) for mortality (22% risk reduction).
- **Manual/non-electronic alerts:** No significant mortality reduction (RR 1.05, 95% CI 0.82–1.35).
- **Bundle adherence improvements:** All alert types improved time-to-blood-culture (RR 1.14, 95% CI 1.03–1.27), time-to-antibiotic, and lactate measurement.
- **Hospital length of stay:** Reduced by 1–2 days across alert system types.

Mechanism of benefit: Alert systems did not fundamentally change sepsis epidemiology but rather improved:

1. **Awareness and mindfulness:** Reminders prompt evaluation of at-risk patients.
2. **Bundle compliance:** Structured order sets ensure protocol adherence.
3. **Time-to-treatment:** Electronic systems reduce delays to antibiotics and fluids.

Limitations and harms:

- **Alert fatigue:** High-sensitivity alerts trigger false positives, leading to alert suppression or cognitive override by clinicians. Few studies quantify this.
- **Overtreatment:** Overly sensitive alerts may result in unnecessary antibiotics in patients with noninfectious inflammation (e.g., post-operative fever, autoimmunity).
- **EHR-dependence:** Systems require robust EHR integration; some hospitals lack mature systems.
- **Equity concerns:** ML-based alerts trained on predominantly high-income, Western datasets may perform poorly in diverse populations or LMICs.

Machine-Learning and Data-Driven Tools

Emerging ML-based sepsis prediction systems (e.g., qSepsis, S-S.M.A.R.T., proprietary hospital algorithms) leverage routine electronic vital signs and lab data to predict sepsis with AUCs often exceeding 0.85. A 2024 Lancet eClinical Medicine study demonstrated a non-laboratory sepsis screening tool achieving AUC 0.84 for sepsis prediction using only triage vitals. However:

- **External validity concerns:** Most models are single-center or regionally trained and show performance degradation when applied to external cohorts (a phenomenon called "model drift").
- **Black-box problem:** Complex ML models lack interpretability, complicating clinical integration and regulatory approval.
- **Fairness and bias:** Algorithms may perpetuate or amplify disparities if training data are non-representative.
- **Cost and infrastructure:** ML systems require substantial IT investment and ongoing model maintenance.

5. Critical Appraisal of Existing Literature: Strengths, Limitations, and Controversies Strengths of the Evidence Base

1. **Large multicenter meta-analyses** (e.g., Wang 2022: 62,338 patients from 26 studies) provide robust pooled estimates of sensitivity, specificity, and diagnostic odds ratios.
2. **Sepsis-3 consensus** (Singer et al., 2016) standardized definitions, improving comparability across studies published after 2016.
3. **Increased prospective studies:** More recent cohorts use prospective enrollment and real-time data capture, reducing retrospective bias.
4. **Diverse settings:** Literature now includes EDs, acute medical/surgical wards, ICUs, and prehospital care, allowing setting-specific synthesis.
5. **Outcome heterogeneity captured:** Meta-analyses report mortality at multiple timepoints (28-day, in-hospital), sepsis bundle adherence, and LOS, acknowledging different clinical endpoints.

Significant Limitations and Controversies

1. Heterogeneity in sepsis definitions:

- Some studies use Sepsis-2 (SIRS + infection), others Sepsis-3 (organ dysfunction). This foundational difference yields incomparable sensitivity/specificity estimates.
- Some use physician-adjudicated sepsis (subjective), others administrative coding, and others prospective SOFA scoring.
- This heterogeneity partially explains the high I^2 (92–99%) in meta-analyses.
- Biological and clinical heterogeneity of sepsis, including distinct host-response phenotypes with differing trajectories and outcomes, further complicates the performance and generalizability of uniform screening tools across populations18

2. The qSOFA controversy:

- Meta-analysis (Wang 2022) shows qSOFA superiority for mortality prediction (DOR 3.79 vs NEWS 2.96 vs SIRS 1.0).
- Yet ED studies (Oi 2025, Wang 2025) show qSOFA's low screening sensitivity (46–66.7%), meaning it misses many early sepsis cases.
- The disconnect arises because of the overlap between mortality prediction ≠ and sepsis diagnosis. qSOFA identifies the sickest (late-stage) patients but is inadequate for early recognition.
- 2021 SSC Guidelines resolved this by recommending against qSOFA as the sole screening tool, while acknowledging its prognostic utility.

3. Reference standard heterogeneity:

- SIRS (≥2 criteria) remains overused as a "sepsis surrogate" despite Sepsis-3.
- SOFA baseline definitions vary: some use baseline SOFA as 0 (assume normal organ function), others use prior ICU admission data.
- This affects the apparent diagnostic accuracy of screening tools.

4. Spectrum and verification bias:

- Many validation studies are single-center, hospital-affiliated, with high sepsis prevalence(e.g., 21–27% of ED cohorts). Real-world ED sepsis prevalence is 3–5%, meaning test characteristics will differ (lower positive predictive value in low-prevalence settings).
- Prospective enrollment bias: Studies often enroll only patients who had blood cultures drawn, excluding lower-risk patients who did not undergo testing.

5. Setting-specific performance:

- qSOFA and SOFA were derived and validated in ICU populations(critically ill). Performance in non-ICU settings (wards, EDs) often differs, particularly when baseline organ function is unknown.
- NEWS/MEWS were designed for general hospital patients, not specifically for sepsis. They capture deterioration broadly, not sepsis specifically.

6. Lactate and biomarker availability:

- Combination tools (qSOFA + lactate) require point-of-care testing or blood gas analysis—unavailable in many EDs in LMICs.
- Studies comparing lactate-inclusive vs -exclusive tools often show inflated accuracy of lactate-inclusive strategies, but generalizability to non-laboratory-rich settings is limited.

7. Dynamic vs. static scoring:

- Most literature evaluates single time-point scores (upon ED arrival). Emerging evidence suggests serial or repeated scores may better capture disease trajectory, but few interventional studies exist.

8. Alert system heterogeneity and unintended harms:

- Electronic alert effectiveness varied substantially (RR 0.78 to 1.05 across studies). Variation driven by alert sensitivity (high sensitivity → alert fatigue → override; low sensitivity → missed cases).
- Few studies measure over treatment harms (unnecessary antibiotics in noninfectious patients, increased antibiotic resistance, healthcare costs).

6. Future Directions and Research Gaps

Critical gaps requiring research:

1. **Prospective, multicenter trials comparing sepsis screening strategies:** A large, pragmatic, cluster-randomized trial comparing NEWS-based screening vs. qSOFA + SOFA vs. local triage score vs. electronic alert, each embedded in quality-improvement bundles, across diverse hospitals (EDs, wards, ICUs) with standardized Sepsis-3 outcomes. Such a trial is notably absent.
2. **Equity and LMIC validation:** Sepsis screening tool validation in diverse populations and low-resource settings is scant. Most evidence comes from high-income, Western, predominantly White cohorts. Multi-country validation studies across diverse microbiology (tropical infections, TB-associated sepsis), comorbidities, and health-system capacity are needed.
3. **Dynamic/serial scoring vs. single time-point:** Prospective studies examining whether repeated sepsis scores(e.g., qSOFA at 1, 3, 6 hours) or sepsis trajectory(rising SOFA) improve diagnostic accuracy and clinical outcomes compared to static assessment.
4. **Objective harm quantification from alert systems:** Randomized trials specifically measuring over treatment burden(unnecessary antibiotic courses, adverse effects, costs, development of multidrug-resistant organisms) alongside mortality and bundle adherence benefits in EHR-alert systems.
5. **Fairness and bias in machine-learning models:** External validation of ML sepsis prediction tools across demographically, geographically, and clinically diverse cohorts; explicit assessment of algorithmic fairness and disparities; development of governance frameworks to monitor model performance in real-world deployment.
6. **Cost-effectiveness analyses:** Comparative economic analyses of different screening strategies (generic score vs. locally optimized tool vs. ML-based alert) considering implementation costs, training, maintenance, and outcomes in different resource settings.
7. **Long-term sepsis outcomes and recovery pathways:** While 2021 SSC guidelines address post-discharge follow-up, evidence on optimal rehabilitation, cognitive screening, and support for sepsis survivors remains nascent.

CONCLUSION

Clinical Synthesis and Practice Recommendations

Key Takeaways

1. **No single tool is universally optimal.**6,8,11 Each screening or prognostic tool trades off sensitivity and specificity. qSOFA is highly specific but poorly sensitive for early detection; SIRS is overly sensitive with inadequate specificity; NEWS/MEWS provide balanced performance but are not sepsis-specific.
2. **Sepsis-3 paradigm shift was correct in emphasizing organ dysfunction**, but implementation in acute settings requires pragmatic, multimodal approaches. The 2021 SSC guideline's strong recommendation

against qSOFA as sole screening tool reflects 5 years of validation data demonstrating its insufficient sensitivity.

3. **Combination strategies outperform single tools.** Evidence (Oi 2025, others) shows that pairing a bedside score (qSOFA or NEWS) with lactate or perfusion assessment achieves both high sensitivity (>85%) and specificity (>80%), addressing the sensitivity-specificity trade-off.
4. **Electronic alert systems reduce mortality and improve bundle adherence**, particularly when rule-based or ML-based alerts are embedded in structured sepsis pathways with order sets and clinician education. Manual alerts without workflow integration show minimal benefit.
5. **Context and setting shape performance.** Tools validated in ICU cohorts (qSOFA, SOFA) perform differently in ED and ward populations. Generic scores (NEWS) must be adapted to local prevalence, infrastructure, and workflow.

Recommendations for Clinical Practice

For Emergency Departments and Acute-Care Settings

Recommended screening approach:

1. **Perform universal vital-sign assessment** at triage/ED arrival using NEWS or MEWS (or institutional adaptation). Flag patients with NEWS ≥5 or MEWS ≥3 as at-risk.
2. **Do NOT rely on qSOFA alone**. If qSOFA ≥2, the patient is at high mortality risk, but qSOFA <2 does NOT exclude sepsis. Use qSOFA in combination with other assessments.
3. **Measure serum lactate** in all patients with suspected infection and vital-sign abnormalities (tachycardia, tachypnea, hypotension). Lactate >2 mmol/L indicates tissue hypoperfusion and warrants urgent sepsis evaluation.
4. **Assess perfusion dynamically**: Skin color/warmth, capillary refill time (normal <2 sec), urine output, mental status. Prolonged CRT or mottling in the context of infection → sepsis evaluation.
5. **Embed screening into workflows**: Use structured sepsis order sets and checklists (e.g., ED sepsis protocol) that automatically trigger assessment, blood cultures, antibiotics, and fluids when NEWS + lactate + clinical suspicion align.
6. **Implement EHR-integrated alerts** if feasible, calibrated to local ED sepsis prevalence and case-mix to minimize alert fatigue. Pair alerts with education and real-time feedback to prevent override.

What NOT to do:

- Do NOT use SIRS criteria alone for sepsis screening (too nonspecific).
- Do NOT defer antibiotic initiation based on qSOFA <2 if clinical suspicion and lactate elevation are present.
- Do NOT assume NEWS ≥5 = sepsis; it indicates risk of deterioration (possibly from non-infectious causes). Integrate the clinical history of infection.

For Non-ICU Hospital Wards

- Implement early warning systems (NEWS/MEWS) for general clinical monitoring.
- Use SOFA scoring for suspected infection to assess organ dysfunction. ΔSOFA ≥2 from admission (or assumed baseline) indicates sepsis.
- Link early warning escalation to infectious disease or critical-care consultation triggers.
- Measure lactate in febrile or unstable patients with suspected infection.

For ICUs

- Use SOFA scoring as part of sepsis diagnosis alongside clinical assessment and microbiologic data.
- qSOFA remains useful for prognostication of mortality in ICU sepsis patients.
- Continue lactate-guided resuscitation (2021 SSC recommendation: target lactate decrease, especially in high initial lactate).

For Resource-Limited Settings (LMICs)

- Prioritize simple, bedside-assessable tools: qSOFA (3 variables) or a simplified version of NEWS (temperature, RR, HR, SBP, consciousness).
- Use qualitative perfusion assessment (skin warmth, capillary refill, urine output) as proxies for lactate when point-of-care testing unavailable.
- Emphasize early antibiotics and fluids (1-hour bundle) over sophisticated diagnostic refinement. The cost of delay exceeds the cost of overtreatment in resource-limited contexts.
- Develop or adapt local triage protocols using available data (vital signs, basic lab tests) rather than adopting generic high-income country tools unchanged.

Limitations of This Review

1. This narrative review does not employ systematic search or meta-analytic synthesis methods, though it references major meta-analyses. A formal PRISMA systematic review would further strengthen recommendations.
2. Heterogeneity remains substantial in the literature due to varying definitions, outcomes, and settings. Complete harmonization is unlikely.
3. Real-world implementation data are limited. Most effectiveness studies are observational; few randomized trials of screening strategies in acute-care exist.
4. Long-term sepsis outcomes and quality-of-life measures are underreported; this review focuses on acute mortality and early bundle adherence.

Declarations

Conflict of Interest

The author(s) declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Funding Statement

- This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

- Author Contributions (ICMJE Format)
- Conceptualization and Design: [Author A, Author B]
- Literature Search and Data Extraction: [Author A, Author C]
- Analysis and Interpretation: [All authors]
- Drafting of Manuscript: [Author A]
- Critical Revision of Manuscript for Important Intellectual Content: [Author B, Author C]

- Final Approval of the Version to be Published: [All authors]
- Agreement to be Accountable for All Aspects of the Work: [All authors]

Ethical Approval

Not applicable. This narrative review synthesizes published data only and does not involve primary human or animal research.

REFERENCES

1. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3).JAMA. 2016;315(8):801–810. doi: 10.1001/jama.2016.0287.
2. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. Lancet. 2020;395(10219):200–211. doi: 10.1016/S0140-6736(19)32989-7.
3. Prescott HC, Iwashyna TJ. Enhancing recovery from sepsis. JAMA. 2018;319(1):62–75. doi: 10.1001/jama.2017.17687.
4. 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. doi: 10.1056/NEJMoa1703058.
5. Liu VX, Fielding-Singh V, Greene JD, et al. The timing of early antibiotics and hospital mortality in sepsis. Am J Respir Crit Care Med. 2017;196(7):856–863. doi: 10.1164/rccm.201609-1848OC.
6. Serafim R, Gomes JA, Salluh J, Povoa P. A comparison of the quick-SOFA and systemic inflammatory response syndrome criteria for the diagnosis of sepsis and prediction of mortality: a systematic review and meta-analysis. Chest. 2018;153(3):646–655. doi: 10.1016/j.chest.2017.12.015.
7. Wang C, Xu R, Zeng Y, et al. A comparison of qSOFA, SIRS, and NEWS in predicting the accuracy of mortality in patients with suspected sepsis: A meta-analysis. PLoS ONE. 2022;17(4):e0266755. doi: 10.1371/journal.pone.0266755.
8. Singer M, Shankar-Hari M. qSOFA, cue confusion. Ann Intern Med. 2018;168(4):293–295. doi: 10.7326/M17-3415.
9. Evans L, Rhodes A, Alhazzani W, et al. Surviving Sepsis Campaign: International Guidelines for the Management of Sepsis and Septic Shock 2021. Crit Care Med. 2021;49(11):e1063–e1143. doi: 10.1097/CCM.0000000000005337.
10. Levy MM, Fink MP, Marshall JC, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Crit Care Med. 2003;31(4):1250–1256. doi: 10.1097/01.CCM.0000050454.01978.3B.
11. Chua WL, Legido-Quigley H, Tee A, et al. Early warning scores for sepsis identification and mortality prediction: A systematic review. J Clin Nurs. 2024;33(10–11):4145–4162. doi: 10.1111/jocn.17061.
12. Oi Y, Ogawa F, Honzawa H, et al. A screening tool to predict sepsis in patients with suspected infection in the emergency department. Cureus. 2025;17(2):e78728. doi: 10.7759/cureus.78728.
13. Puskarich MA, Shapiro NI, Trzeciak S, et al. Prognostic value and physiologic correlates of serum lactate concentrations in early sepsis. J Crit Care. 2012;27(4):432–438. doi: 10.1016/j.jcrc.2012.02.009.
14. Filbin MR, Wolf SJ, Lipman GS, et al. Sepsis Alert Systems, Mortality, and Adherence in Emergency Departments: A Systematic Review and Meta-Analysis. JAMA Netw Open. 2024;7(6):e2821277. doi: 10.1001/jamanetworkopen.2024.21277.
15. Ferrari R. Writing narrative style literature reviews. Med Writ. 2015;24(4):230–235.
16. Wang W, Wang K, Wang Y, et al. Comparison and enhancement of sepsis prediction models in emergency care: Insights from CETAT and MIMIC-IV databases. Biomol Biomed. 2025;25(7):1470–1478. doi: 10.17305/bb.2024.11134.

17. Jayaprakash N, O'Neal HR Jr, Dukes SF. Reassessing shock index and modified shock index in the setting of sepsis. Am J Emerg Med. 2015;33(3):448–449. doi: 10.1016/j.ajem.2014.12.019.

18. Seymour CW, Kennedy JN, Wang S, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for patients with sepsis. JAMA. 2019;321(20):2003–2017. doi: 10.1001/jama.2019.5791.

qSOFA	**Wang 2022 (62,338)**	**Mixed**	**Mortality prediction**	**DOR 3.79 (highest specificity 82%)**	**High prognostic accuracy; low sensitivity (46%)**
SIRS	Wang 2022	Mixed	Mortality prediction	DOR 1.0 (sensitivity 82%, specificity 24%)	Overly sensitive; poor specificity
NEWS	Wang 2022	Mixed	Mortality prediction	DOR 2.96 (intermediate)	Balanced; alternative to qSOFA
qSOFA	Oi 2025 (75)	ED	Sepsis (organ dysfunction)	AUC 0.833, Sens 66.7%, Spec 100%	Excellent specificity; poor screening sensitivity
NEWS	Oi 2025	ED	Sepsis	AUC 0.846, Sens 69%, Spec 93%	Balanced performance; superior to qSOFA
MEWS	Oi 2025	ED	Sepsis	AUC 0.846, Sens 75%, Spec 93%	Equivalent to NEWS; slightly higher sensitivity
NEWS	Wang 2025 (12,799)	ED	Sepsis	AUC 0.737 (sepsis); 0.653 (high-risk)	Outperforms qSOFA (0.64) and SIRS (0.71)
qSOFA + Lactate	Oi 2025	ED	Sepsis	AUC 0.844, Sens 87.5%, Spec 81.4%	Combination achieves >80% sensitivity AND specificity

Supplementary Tables

Table 1. Comparative Diagnostic Accuracy of Sepsis Screening Tools Across Settings

Table 2. Electronic and Manual Sepsis Alert System Effectiveness (Meta-Analysis, 2024; 22 studies, 19,580 patients)

Any Alert System	**22**	**0.83 (0.72–0.96)**	**~1–2 days**	**Improved (RR 1.10–1.14)**	**17% mortality reduction**
Electronic/EHR-Integrated	~15	**0.78 (0.67–0.92)**	1.5–2 days	**RR 1.14 (blood culture)**	Superior to manual; 22% mortality reduction
Manual/Notification-Only	~7	1.05 (0.82–1.35)	Minimal	Limited improvement	No significant mortality benefit
ML-Based	Subset	~0.75–0.80	2–3 days	Excellent	Emerging evidence; external validation limited

Table 3. Recommended Sepsis Screening Strategies by Clinical Setting

ED (all patients)	**NEWS ≥5 or MEWS ≥3**	**qSOFA + Lactate**	**Antibiotics <1 h (shock), <3 h (no shock)**	**Structured sepsis order set; EHR alert integration**
ED (low lactate access)	qSOFA + Perfusion assessment (CRT, skin turgor)	SOFA organ assessment	Antibiotics <1 h	Bedside tools only
Acute medical ward	SIRS ≥2 or NEWS ≥5	SOFA score, lactate	Clinical team assessment <3 h	Escalation protocol to ICU/ED if ΔSOFA ≥2
ICU	SOFA (≥2 from baseline)	qSOFA for prognostication	Sepsis bundle adherence monitored	Lactate-guided resuscitation
LMIC setting	qSOFA or simplified NEWS (5 vitals)	Lactate if available; perfusion signs	Protocol-driven bundle; avoid "diagnosis-delay"	Emphasis on early antibiotics over diagnostic refinement