Journal of Pioneering Medical Sciences
Received: March 22, 2026 | Accepted: September 19, 2026 | Published: November 05, 2026
Volume 15, Issue 10, Pages 199-205

DOI https://doi.org/10.47310/jpms2026151020



Artificial Intelligence for Early Sepsis Recognition and Timely Antibiotic Administration across Emergency and Critical Care Settings: A Systematic Review of Diagnostic Performance and Clinical Outcomes

Moamen Abdelfadil Ismail[1*], Abid Waseem[2], Salem Yousef Bamuawwadhah[3], Ahmed Aamir Butt[4], Reem Abdullah Alrashed[5], Mohamed Magdy Ali Allam[6], Mehnaz Iftikhar[7], Syed Muhammad Zeeshan Ahmed[8], Imtiaz Ahmed Soomro[9], Abdulrahman Adnan B. Alzanbagi[10], Murad Mamdouh Aleidiney[11], Aeraj Mohammed Abdul Samad Peerzade[12], Abdullah Malik Zaheer[13] and Syed Zulkhar[14]

[1]Internal 1King Abdulaziz Specialist Hospital, Sakaka, Aljouf, Saudi Arabia
[2]Emergency Medicine, National Guard Hospital, Madina, Saudi Arabia
[3]Intensive Care Unit, Dr. Soliman Fakeeh Hospital, Jeddah, Saudi Arabia
[4]National Guard Health Affairs, Emergency Medicine, Saudi Arabia
[5]Batterjee Medical College, Saudi Arabia
[6]Internal Medicine Department, King Abdulaziz Hospital- National Guard Health Affairs, Saudi Arabia
[7]Almana General Hospital, Saudi Arabia
[8]Intensive Care Unit Physician, Almoosa Specialist Hospital, Saudi Arabia
[9]Staff Physician Emergency Physician, National Guard Hospital Madina, Saudi Arabia
[10]Medical Student
[11]Medical Student
[12]Johns Hopkins Aramco Healthcare, Saudi Arabia
[13]Emergency Medicine, John Hopkins Aramco Health Care, Saudi Arabia
[14]Emergency Medicine, National Guards Health Affairs, National Gaurds Health Affairs, Saudi Arabiaa

Author Designation: [11]Medical Student

*Corresponding author: Moamen Abdelfadil Ismail (e-mail: moamen.fadil83@gmail.com).

©2026 the Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0

Abstract **Objectives:** Sepsis remains a leading cause of morbidity and mortality, where early recognition and timely antibiotic administration are critical. Artificial Intelligence (AI) has been increasingly investigated for sepsis prediction and clinical decision support across emergency and critical care settings. **Objective:** To systematically evaluate the diagnostic performance of AI models for early sepsis recognition and their reported effects on time-to-antibiotic administration and selected clinical outcomes. **Methods:** Following PRISMA 2020 guidelines, we searched PubMed, Scopus, Web of Science, Embase and Google Scholar from 2015 to January 2026. Studies evaluating AI-based models for sepsis detection that reported diagnostic accuracy, time-to-antibiotics, mortality, or length of stay were included. Two reviewers independently screened and extracted data. **Results:** Ten studies were included (randomized trials, prospective and retrospective cohorts and before-and-after studies) conducted in emergency departments, intensive care units, hospital wards and prehospital settings. AI models showed AUROC values ranging from 0.83 to 0.92 across different validation approaches. Several implementation studies reported reductions in time-to-antibiotic administration (absolute differences ranging from approximately 10-16 minutes, or higher proportions of patients receiving antibiotics within 1-3 hours). Effects on mortality and length of stay were inconsistent across studies. **Conclusion:** AI-based systems demonstrate promising diagnostic performance for early sepsis recognition and, in selected implementation studies, appear to accelerate antibiotic administration. However, substantial heterogeneity in study designs, settings, sepsis definitions and outcome measures limits the ability to draw firm conclusions about consistent clinical effectiveness. Further prospective validation and implementation research are needed.

Key Words Sepsis, Artificial Intelligence, Machine Learning, Systematic Review, Early Detection, Clinical Decision Support Systems, Time-to-Antibiotics, Antimicrobial Therapy, Emergency Department, Early Warning Systems

INTRODUCTION

Sepsis remains one of the most serious global health challenges, defined as life-threatening organ dysfunction caused by a dysregulated host response to infection. Despite continuous advances in critical care, sepsis continues to be associated with high rates of morbidity, mortality and healthcare expenditure. Timely recognition and early administration of appropriate antibiotics are among the most important factors that influence patient outcomes. Nevertheless, delays in diagnosis and treatment still occur frequently, especially in busy clinical environments such as emergency departments, where patients often present with non-specific or overlapping symptoms [1].

For many years, clinicians have relied on traditional scoring systems such as the Systemic Inflammatory Response Syndrome (SIRS) criteria, the Sequential Organ Failure Assessment (SOFA) score and the quick SOFA (qSOFA) score. Although these tools are widely used, they have important limitations, including relatively low sensitivity and specificity and limited ability to provide continuous, real-time risk assessment. These shortcomings have stimulated growing interest in Artificial Intelligence (AI) and machine learning approaches that can process large volumes of electronic health record data and detect subtle patterns that may precede clinical deterioration.

In the context of this review, "early recognition" is defined as the capacity of an AI system either to predict sepsis before it becomes clinically obvious or to generate an alert that enables clinicians to assess and treat patients earlier than would occur under usual care. It is essential to distinguish between two related but different concepts. The first is the diagnostic or predictive performance of an AI model, which is commonly measured by metrics such as the area under the receiver operating characteristic curve. The second is the actual clinical effectiveness of an implemented AI alert system in changing clinician behaviour, reducing the time to antibiotic administration and improving patient outcomes. High predictive accuracy alone does not automatically guarantee better clinical processes or better patient results.

Earlier systematic reviews and meta-analyses have concentrated mainly on the diagnostic accuracy of AI models for sepsis prediction. Reviews such as those by Fleuren *et al.* [2] and Yadgarov *et al.* [3] have summarised model performance but have given limited attention to downstream clinical effects, particularly the impact on antibiotic timing and patient-centred outcomes in real-world settings. The present review differs from previous work by examining both diagnostic performance and the reported effects of AI systems on time-to-antibiotic administration and selected clinical outcomes. In addition, it includes studies conducted in emergency departments, intensive care units, hospital wards and prehospital settings. These different environments were retained because early recognition of sepsis and timely antibiotic delivery are clinically relevant across the acute care continuum and because limiting the review solely to emergency department populations would exclude a substantial body of implementation evidence. Setting-related differences are considered when interpreting the findings.

Saudi Arabia's Vision 2030 places strong emphasis on digital transformation of the healthcare system, including the use of AI-enabled tools to improve the quality and efficiency of emergency and critical care. While this national context provides relevant motivation, the primary scientific rationale for the present review is the identified gap in evidence concerning the clinical effectiveness of AI systems for sepsis, rather than policy considerations alone.

The objective of this systematic review is to evaluate the diagnostic performance of AI-based models for early sepsis recognition and to assess the reported effects of these systems on time-to-antibiotic administration and selected patient outcomes, specifically mortality and length of stay, across emergency and critical care settings.

METHODS

Study Design and Review Protocol

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. The primary objective was to evaluate the diagnostic performance of Artificial Intelligence (AI) based models for early sepsis recognition and to assess their reported effects on time-to-antibiotic administration and selected clinical outcomes across emergency and critical care settings.

A review protocol was not prospectively registered in PROSPERO or any other public repository. The eligibility criteria, primary outcomes and data-synthesis approach were defined before the formal screening of studies began. The absence of prospective registration is acknowledged as a methodological limitation.

Eligibility Criteria

Inclusion criteria were as follows:

- **Population:** Adult or pediatric patients with suspected or confirmed sepsis in Emergency Departments (EDs), Intensive Care Units (ICUs), hospital wards, or prehospital settings
- **Intervention/Exposure:** Use of AI-based systems, including machine learning algorithms, deep learning models, or AI-powered early warning/alert systems for sepsis prediction or detection. Both rule-based systems enhanced by AI and purely data-driven models were eligible
- **Comparators:** Usual care, conventional clinical scoring systems (e.g., SIRS, qSOFA, SOFA, MEWS), or pre-implementation periods
- **Outcomes:** Diagnostic performance metrics (AUROC, sensitivity, specificity, or related measures), time-to-antibiotic administration, mortality, length of stay, or organ dysfunction
- **Study Designs:** Randomized controlled trials, prospective or retrospective cohort studies, before-and-after studies and model validation studies that reported patient-level or clinical outcomes
- **Language and Period:** English-language full-text articles published between 2015 and January 2026

Exclusion criteria included:

- Non-empirical publications (editorials, commentaries, narrative reviews, conference abstracts without full text)
- Studies that developed AI models without any evaluation of diagnostic performance or clinical outcomes
- Studies focused solely on general clinical deterioration without sepsis-specific outcomes
- Duplicate publications or overlapping cohorts (the most complete or recent report was retained)

The decision to include studies from both ED and ICU (and related) settings was made because early sepsis recognition and timely antibiotic administration are clinically relevant across the acute care continuum. Findings are interpreted with attention to setting-related differences.

Search Strategy

A comprehensive literature search was performed in PubMed, Scopus, Web of Science, Embase and Google Scholar. The final search was completed in January 2026. The search combined controlled vocabulary and free-text terms related to sepsis, artificial intelligence/machine learning, early detection/prediction, emergency or critical care settings and antibiotic timing or clinical outcomes.

The core Boolean search string included combinations of the following terms: ("sepsis" OR "septic shock") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "early warning system" OR "clinical decision support") AND ("early detection" OR "prediction" OR "recognition" OR "alert") AND ("emergency department" OR "ICU" OR "intensive care" OR "prehospital") AND ("time to antibiotics" OR "antibiotic timing" OR "mortality" OR "length of stay").

Reference lists of included studies and relevant systematic reviews were manually screened for additional eligible articles. All retrieved records were imported into Zotero and duplicates were removed both automatically and by manual verification.

The English-language restriction is acknowledged as a potential source of language bias. Full database-specific search strategies are available upon request to support reproducibility.

Study Selection Process

Two reviewers independently screened titles and abstracts against the eligibility criteria. Full-text articles of potentially relevant records were then assessed independently by the same reviewers. Disagreements were resolved through discussion; if consensus could not be reached, a third reviewer was consulted. The selection process is summarized in a PRISMA flow diagram (Figure 1), which reports the number of records identified, duplicates removed, records screened, full-text articles assessed, reasons for exclusion and studies included in the final review.

Data Extraction

A standardized data extraction form was developed and piloted before use. Two reviewers independently extracted data and a third reviewer checked the extracted information for accuracy and completeness. The following variables were extracted:

- Study characteristics (authors, year, country, design, setting)
- Population details (sample size, age group, sepsis definition)
- AI model description (type of algorithm, input features, validation approach, implementation status)
- Comparator
- Diagnostic performance metrics (AUROC, sensitivity, specificity and related measures where reported)
- Time-to-antibiotic outcomes (definition of the time origin, summary statistic, absolute or relative differences)
- Clinical outcomes (mortality definition and follow-up period, length of stay, organ dysfunction)
- Key limitations reported by the original authors

Missing or unclear information was recorded as such; study authors were not contacted for additional data.

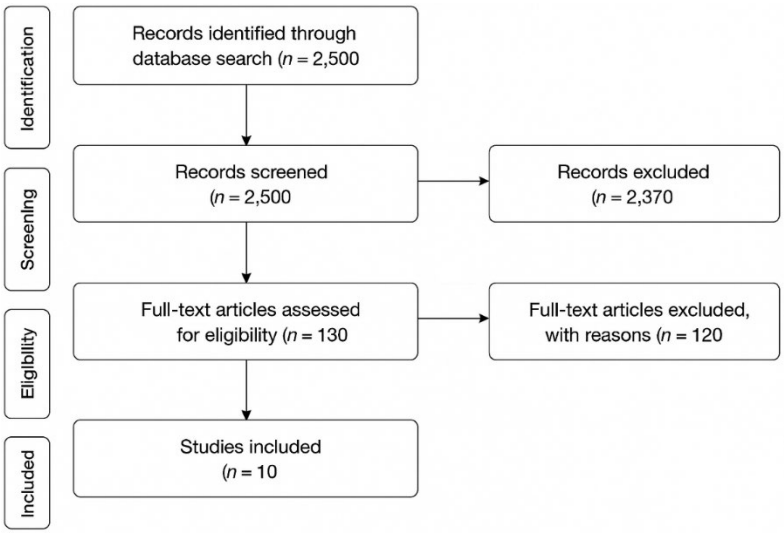


Figure 1: PRISMA Flow Diagram

Risk-of-Bias Assessment

Risk of bias was assessed according to study design. The Cochrane Risk of Bias 2 (RoB 2) tool was used for randomized trials. The Newcastle-Ottawa Scale (NOS) was applied to observational cohort and before-and-after studies. For studies whose primary focus was diagnostic accuracy, the relevance of a diagnostic-specific tool (such as QUADAS-2) was considered. Judgments were made at the domain level and summarized as low, moderate, or high risk of bias. Risk-of-bias assessments informed the interpretation of findings rather than being used as formal inclusion/exclusion criteria.

Data Synthesis

Due to substantial clinical and methodological heterogeneity in study designs, populations, AI systems, sepsis definitions, prediction horizons, antibiotic-timing endpoints and outcome measures, a quantitative meta-analysis was not performed. Results were synthesized narratively and organized into the following domains:

- Diagnostic performance of AI models
- Effects on time-to-antibiotic administration
- Effects on mortality and length of stay
- Implementation considerations and real-world applicability

Findings from model development/validation studies were distinguished from those of clinical implementation studies. Quantitative results (effect estimates, confidence intervals and p-values) are reported where available in the original studies.

Ethical Considerations

This systematic review used only previously published data and did not involve human participants or primary data collection; therefore, ethical approval and informed consent were not required. All included studies were assumed to have obtained appropriate institutional ethical approval.

RESULTS

Characteristics of Included Studies

Ten studies met the inclusion criteria and were included in this review. The studies employed a range of designs, including randomized controlled trials, prospective and retrospective cohort studies and before-and-after evaluations. Sample sizes varied widely, from small single-center cohorts (n = 142) to large multi-center datasets exceeding 500,000 patients. Clinical settings included adult emergency departments, pediatric emergency departments, intensive care units, hospital wards and prehospital emergency medical services. The studies were conducted in the United States, Thailand, India, China, Switzerland and Germany.

The artificial intelligence approaches evaluated included machine learning-based early warning systems, random forest models and proprietary clinical decision support tools (such as TREWS, AISE and HERACLES). Some studies assessed model performance in retrospective or validation cohorts, while others evaluated the impact of implemented AI alerts on clinical processes and patient outcomes. A summary of study characteristics is presented in Table 1.

Diagnostic Performance

AI models demonstrated AUROC values ranging from 0.83 to 0.92 across the included studies. Nemati *et al.* [4] reported AUROC values of 0.83-0.85 for predicting sepsis up to 12 hours in advance in an ICU cohort.

Table 1: Characteristics of the Included Studies

Study	Country	Design	Setting	Sample Size	AI Model/System	Comparator	Key Outcomes	Main Findings
Kijpaisalratana *et al.* [7]	Thailand	Cluster RCT	ED	574	ML alert system	Standard alerts	TTA, diagnostic accuracy	Higher % received antibiotics ≤1 h (68.4% Vs 60.1%) and ≤3 h; superior AUROC Vs qSOFA/SIRS/MEWS
Adams *et al.* [11]	USA	Prospective cohort	Multi-center (hospital-wide)	590,736	TREWS	Usual care	Mortality, LOS, organ failure	Early alert confirmation associated with lower adjusted mortality and shorter LOS
Shimabukuro *et al.* [10]	USA	RCT	ICU	142	Machine learning algorithm	Standard care	Mortality, LOS	In-hospital mortality reduced (40.0% vs 13.6% in sepsis subgroup); LOS reduced (13.0 to 10.3 days)
Mixon *et al.* [8]	USA	Retrospective	ED / EMS	507	Sepsis alert timing	ED Vs prehospital alert	TTA	Median TTA 48.5 Vs 64.5 min ($p<0.001$); higher % within 60 min
Joshi *et al.* [6]	India	Observational	ED	2,000	AI risk score	SIRS/ qSOFA	Diagnostic accuracy	AUROC 0.92; sensitivity 88%, specificity 85%
Kandaswamy *et al.* [9]	USA	Retrospective (before-after)	Pediatric ED	599,163	Vendor AI sepsis model	Pre-implementation	TTA, fluid bolus	Time to first antibiotic reduced by 10 min; time to fluid bolus reduced by 16.7 min
Despraz *et al.* [12]	Switzerland	Before-after	Hospital wards	123,410	HERACLES AI system	Control wards	Mortality	Reduced in-hospital and 90-day mortality in intervention wards
Wang *et al.* [5]	China	Cohort	ICU	4,449	Random forest	Clinical assessment	Prediction performance	AUROC 0.91
Nemati *et al.* [4]	USA	Cohort	ICU	>70,000	AISE	Clinical assessment	Prediction performance	AUROC 0.83–0.85 (up to 12 h before onset)
Pohlan *et al.* [13]	Germany	Retrospective	ED	357	CT imaging (septic foci)	Clinical assessment	Source detection	Positive predictive value 81.8%

Wang *et al.* [5] achieved an AUROC of 0.91 using a random forest model in ICU patients. Joshi *et al.* [6] reported the highest AUROC of 0.92 in an emergency department cohort, with sensitivity of 88% and specificity of 85%, outperforming SIRS (AUROC 0.78) and qSOFA (AUROC 0.74). Kijpaisalratana *et al.* [7] also found superior discrimination of their machine learning alert system compared with qSOFA, SIRS and MEWS in an emergency department randomized trial.

These performance estimates were derived under varying conditions, including internal validation, temporal validation and prospective evaluation. Direct head-to-head comparisons with traditional scoring systems were not available in every study and differences in sepsis definitions, prediction horizons and operating thresholds limit direct comparability of the reported AUROC values. Sensitivity, specificity and predictive values were inconsistently reported across studies.

Time-to-Antibiotic Administration

Several implementation studies reported reductions in time-to-antibiotic administration. In a cluster-randomized trial conducted in an emergency department, Kijpaisalratana *et al.* [8] observed that a higher proportion of patients in the AI-alert group received antibiotics within 1 hour (68.4% Vs 60.1%, p = 0.04) and within 3 hours (94.5% Vs 89.0%, p = 0.02) compared with standard alerts.

Mixon *et al.* [8] found that prehospital sepsis alerts were associated with faster antibiotic administration than emergency department-initiated alerts (median 48.5 minutes Vs 64.5 minutes, p<0.001), with a higher proportion of patients receiving antibiotics within 60 minutes (59.1% Vs 44%, p = 0.01).

In a pediatric emergency department before-and-after study, Kandaswamy *et al.* [9] reported a reduction in time to first antibiotic from 112 minutes to 102 minutes (p = 0.05) and a decrease in time to fluid bolus of 16.7 minutes (p = 0.03) after implementation of an AI model.

These findings suggest that AI alerts can improve process measures related to antibiotic timing in selected settings. However, the studies used different starting points for measuring time-to-antibiotics (e.g., ED arrival, alert activation, or sepsis recognition) and not all studies evaluated prospectively implemented systems.

Mortality and Length of Stay

Effects on mortality and length of stay were less consistent. Shimabukuro *et al.* [10], in a small randomized trial conducted in an intensive care unit (n = 142), reported a relative reduction in in-hospital mortality of 58% in the sepsis subgroup (40.0% Vs 13.6%, p = 0.023) and a reduction in hospital length of stay from 13.0 to 10.3 days (p = 0.042).

Adams *et al.* [11], in a large multi-center prospective evaluation of the TREWS system, found that early provider confirmation of AI alerts was associated with an adjusted absolute mortality reduction of 3.3 percentage points, reduced organ failure and shorter hospital stays. Despraz *et al.* [12] observed reductions in in-hospital and 90-day mortality in wards using an AI system, while control wards showed no improvement.

In contrast, Mixon *et al.* [8] and Kandaswamy *et al.* [9] reported no statistically significant differences in mortality despite improvements in process measures. The available evidence therefore does not demonstrate a consistent mortality benefit across study designs and settings. Improvements in predictive performance or antibiotic timing did not uniformly translate into better patient-centered outcomes.

Study Heterogeneity and Implementation Considerations

The included studies differed substantially in AI algorithms, data sources, sepsis definitions, prediction time windows, clinical settings and implementation strategies. Some systems operated in real time and generated clinician-facing alerts integrated into electronic health records, while others were evaluated only retrospectively. Information on clinician training, alert response rates, false-positive burden and workflow impact was limited and inconsistently reported. These differences contribute to the heterogeneity of findings and limit the ability to draw uniform conclusions about clinical effectiveness.

DISCUSSION

This systematic review examined the diagnostic performance of Artificial Intelligence (AI) models for early sepsis recognition and their reported effects on time-to-antibiotic administration and selected clinical outcomes. The included studies show that AI models can achieve good discrimination, with AUROC values commonly reported between 0.83 and 0.92. In several implementation studies, AI-supported alerts were associated with faster antibiotic administration. However, the evidence is heterogeneous in design, setting, population, sepsis definition and outcome measurement and the findings must be interpreted with caution.

It is important to distinguish three levels of evidence. The first is predictive or diagnostic performance (for example, AUROC). The second is change in clinical processes, such as time-to-antibiotic administration. The third is patient-centered outcomes, including mortality and length of stay. High predictive accuracy does not automatically translate into improved clinical decision-making or better patient outcomes. Threshold selection, calibration, disease prevalence, false-positive rates and the degree of integration into clinical workflow all influence whether an accurate model produces meaningful benefit.

Several studies demonstrated reductions in time-to-antibiotic administration after implementation of AI alerts. These process improvements are clinically relevant because earlier antibiotics are associated with better outcomes in sepsis. Nevertheless, the studies used different starting points for measuring timing (for example, emergency department arrival versus alert activation) and some evaluations were retrospective rather than prospective implementations. Before-and-after studies are particularly vulnerable to confounding from concurrent changes in sepsis protocols, staffing, or clinical awareness.

Effects on mortality and length of stay were inconsistent. A small randomized trial and some larger observational evaluations reported lower mortality or shorter stays in association with AI systems or timely response to alerts. Other studies found no significant difference in mortality despite improvements in process measures. The current evidence base does not support a conclusion that AI systems consistently reduce mortality.

Compared with earlier systematic reviews that focused mainly on diagnostic accuracy [2, 3], the present review adds attention to antibiotic timing and clinical outcomes. At the same time, the breadth of settings (emergency departments, intensive care units, wards and prehospital care) and the mix of development, validation and implementation studies limit the ability to draw uniform conclusions specifically for emergency department practice.

Generalizability and Context

Most of the included studies were conducted in relatively well-resourced healthcare systems with mature electronic health record infrastructure. Differences in laboratory turnaround times, staffing models, sepsis protocols and case mix may affect the performance and impact of AI systems when transferred to other settings. Local validation of calibration, sensitivity, specificity, false-positive rates and clinical impact is therefore necessary before widespread adoption. Recommendations for clinical deployment should include prospective evaluation and ongoing monitoring rather than reliance solely on published performance metrics from other institutions.

Saudi Arabia's Vision 2030 emphasizes digital transformation and the use of AI to improve healthcare quality and efficiency. The findings of this review may help inform future local evaluation of AI tools for sepsis, but they do not directly demonstrate effectiveness within Saudi emergency or critical care settings.

Implications for Policy and Practice

For healthcare systems adopting AI-based sepsis detection, key recommendations include:

- Mandatory external validation before clinical deployment
- Integration with electronic health records without increasing alert fatigue
- Continuous performance monitoring across diverse patient populations
- Alignment with national digital health strategies such as Saudi Vision 2030. Residency programs seeking premium recognition should consider including AI training modules and supporting implementation science projects

Limitations

This review has several important limitations. First, substantial heterogeneity in study design, clinical setting, patient population (adult and pediatric), sepsis definitions, prediction horizons, antibiotic-timing endpoints and validation methods limited the ability to synthesize results quantitatively and prevented a formal meta-analysis. Second, the inclusion of studies from emergency departments, intensive care units, hospital wards and prehospital settings reduces the specificity of conclusions for emergency department practice alone. Third, many of the studies evaluating clinical outcomes were observational or used before-and-after designs and are therefore susceptible to confounding. Consequently, the review cannot establish a causal effect of AI systems on mortality. Fourth, the restriction to English-language publications may have introduced language bias. Fifth, the search strategy and screening process, while systematic, may not have captured every relevant study, particularly given the dynamic nature of Google Scholar results. Finally, no formal certainty-of-evidence assessment (for example, GRADE) was performed. These limitations mean that conclusions about clinical effectiveness remain provisional.

CONCLUSION

AI-based models show promising diagnostic performance for early sepsis recognition, with reported AUROC values typically in the range of 0.83-0.92. In selected implementation studies, AI alerts have been associated with reductions in time-to-antibiotic administration. However, the evidence is heterogeneous and improvements in predictive accuracy or process measures have not consistently translated into lower mortality or shorter length of stay. The current body of evidence does not support firm conclusions about widespread clinical effectiveness.

Successful clinical use of these systems requires external validation in the target population, prospective evaluation of workflow impact, monitoring of false-positive alerts and assessment of potential unintended consequences. While the findings may inform digital health initiatives aligned with Saudi Vision 2030, local evaluation remains essential. Future research should prioritize rigorous comparative and implementation studies that clearly link AI alerts to changes in clinician behavior and patient-centered outcomes.

REFERENCES

[1] Fay, K. *et al.* "Assessment of health care exposures and outcomes in adult patients with sepsis and septic shock." *JAMA Network Open*, vol. 3, no. 7, 2020, e206004.

[2] Fleuren, L. M. *et al.* "Machine learning for the prediction of sepsis: A systematic review and meta-analysis of diagnostic test accuracy." *Intensive Care Medicine*, vol. 46, no. 3, 2020, pp. 383-400.

[3] Yadgarov, M.Y. *et al.* "Early detection of sepsis using machine learning algorithms: A systematic review and network meta-analysis." *Frontiers in Medicine*, vol. 11, 2024, article 1491358.

[4] Nemati, S. *et al.* "An interpretable machine learning model for accurate prediction of sepsis in the ICU." *Critical Care Medicine*, vol. 46, no. 4, 2018, pp. 547-553.

[5] Wang, D. *et al.* "A machine learning model for accurate prediction of sepsis in ICU patients." *Frontiers in Public Health*, vol. 9, 2021, article 754348.

[6] Joshi, M.A. *et al.* "Evaluation of artificial intelligence-based risk scoring for early sepsis detection in emergency department admissions." *Journal of Heart Valve Disease*, vol. 30, 2025, pp. 166-169.

[7] Kijpaisalratana, N. *et al.* "Real-time machine learning-assisted sepsis alert enhances the timeliness of antibiotic administration and diagnostic accuracy in emergency department patients with sepsis: A cluster-randomized trial." *Internal and Emergency Medicine*, vol. 19, no. 5, 2024, pp. 1415-1424.

[8] Mixon, M. *et al.* "Time to antibiotic administration: Sepsis alerts called in emergency department versus in the field via emergency medical services." *American Journal of Emergency Medicine*, vol. 44, 2021, pp. 291-295.

[9] Kandaswamy, S. *et al.* "Early clinical evaluation of a vendor-developed pediatric artificial intelligence sepsis model in the emergency department." *Journal of the American Medical Informatics Association*, vol. 32, no. 10, 2025, pp. 1542-1551.

[10] Shimabukuro, D.W. *et al.* "Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: A randomized clinical trial." *BMJ Open Respiratory Research*, vol. 4, no. 1, 2017, article e000234.

[11] Adams, R. *et al.* "Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis." *Nature Medicine*, vol. 28, no. 7, 2022, pp. 1455-1460.

[12] Despraz, J. *et al.* "An artificial intelligence-powered learning health system to improve sepsis detection and quality of care: A before-and-after study." *npj Digital Medicine*, 2026.

[13] Pohlan, J. *et al.* "Computed tomography for detection of septic foci: Retrospective analysis of patients presenting to the emergency department." *Clinical Imaging*, vol. 69, 2021, pp. 223-227.