

Diagnostic Performance and Limitations of Routine Vital Signs and Early Warning Scores in Postpartum Maternal Sepsis: A Systematic Review

Review began 09/22/2026

Review ended 10/17/2026

Published 10/20/2026

© Copyright 2026

Elshaikh et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118268

Reem E. Elshaikh ¹, Thouraya Mohamed Ahmed Elawad ¹, Zainab Alshami ², Haytham Mohammed Osman Altigani ³, Asim Ahmed ⁴

1. Emergency Medicine, Johns Hopkins Aramco Healthcare, Al-Ahsa, SAU 2. Emergency Medicine, Hayat National Hospital, Jazan, SAU 3. Emergency Medicine, Ministry of Health, Riyadh, SAU 4. Medicine and Surgery, University of Gezira, Wad Madani, SDN

Corresponding author: Thouraya Mohamed Ahmed Elawad, thouraya.moh@gmail.com

Abstract

Maternal sepsis after childbirth can be difficult to recognize because expected postpartum physiological changes and common obstetric complications may resemble early infection. Conversely, sepsis may initially occur without marked physiological abnormalities. This systematic review evaluated the diagnostic performance and limitations of routine vital signs, shock index, and vital sign-based screening systems for the early recognition of postpartum maternal sepsis. PubMed/MEDLINE, Scopus, Web of Science Core Collection, and Google Scholar were searched through July 13, 2026, with supplementary citation and manual searching. Eligible studies included postpartum populations and mixed obstetric populations with relevant postpartum data. Two reviewers independently screened studies, extracted data, and assessed risk of bias using design-appropriate tools. The primary synthesis examined diagnostic and screening performance, while a secondary synthesis considered severity prediction. Owing to substantial clinical and methodological heterogeneity, findings were synthesized narratively without meta-analysis. Nineteen studies were included. Higher-sensitivity screening approaches generally produced more alerts and lower specificity, whereas more specific approaches identified a smaller proportion of cases. In mixed obstetric evidence, systemic inflammatory response syndrome favored sensitivity at the expense of specificity, while quick Sequential Organ Failure Assessment showed greater specificity but lower sensitivity. Abnormal temperature was frequently observed without confirmed infection, while sepsis could also occur without fever; other vital signs and composite scores showed similarly context-dependent performance. Shock index was more informative for severity prediction after suspected or confirmed infection than for initial diagnosis. Most studies had important methodological limitations. Current evidence does not support the use of any isolated vital sign or warning score to confirm or exclude postpartum maternal sepsis. Repeated complete observations, assessment of physiological trends, and interpretation within the clinical context may improve recognition.

Categories: Emergency Medicine, Epidemiology/Public Health, Internal Medicine

Keywords: maternal sepsis, postpartum period, puerperal infection, shock index, vital signs

Introduction And Background

Maternal sepsis remains an important and potentially preventable cause of maternal death. A recent WHO systematic analysis estimated that pregnancy-related infection accounted for approximately 7% of maternal deaths globally during 2009 to 2020, with most maternal deaths from sepsis occurring during the postpartum period [1]. National surveillance from the United Kingdom showed that severe maternal sepsis can progress rapidly even within well-resourced maternity services [2]. This underscores the importance of prompt recognition. The postpartum period presents a particular diagnostic challenge because infection may develop after an apparently uncomplicated birth, during recovery from obstetric interventions or complications, or after discharge from maternity care.

Several physiological abnormalities used to identify sepsis may also occur during normal recovery after childbirth. Pain, blood loss, medications, anesthesia, surgery, and fluid shifts may contribute to tachycardia, tachypnea, temperature changes, or hypotension without invasive infection. Conversely, maternal sepsis may initially present without marked physiological abnormalities. The limited specificity of conventional systemic inflammatory response syndrome (SIRS) criteria has also been demonstrated in general adult populations [3].

Early warning systems use physiological observations and predefined thresholds to identify women requiring further assessment. Their interpretation depends on the population, thresholds, and clinical outcomes evaluated. A recent scoping review of obstetric early warning systems found substantial heterogeneity in the index tests and target conditions, limiting direct comparison across studies [4].

How to cite this article

Elshaikh R E, Elawad T, Alshami Z, et al. (October 20, 2026) Diagnostic Performance and Limitations of Routine Vital Signs and Early Warning Scores in Postpartum Maternal Sepsis: A Systematic Review. Cureus 18(10): e118268. DOI 10.7759/cureus.118268

The Sepsis-3 consensus defined sepsis as life-threatening organ dysfunction caused by a dysregulated host response to infection, shifting emphasis away from inflammation alone [5]. Pregnancy-related and postpartum physiological changes can complicate the application of general adult thresholds in maternity care. Obstetric-specific early warning systems were developed to identify physiological deterioration in this context. However, a recent systematic review of maternal early-warning scores identified substantial methodological weaknesses, high risk of bias in most evaluated studies, and limited independent validation [6]. The WHO defined maternal sepsis as a life-threatening condition characterized by organ dysfunction resulting from infection during pregnancy, childbirth, the post-abortion period, or the postpartum period [7]. This definition distinguishes maternal sepsis from infection without organ dysfunction.

Maternal sepsis may arise from obstetric or non-obstetric infections, and recognition requires integration of physiological observations with symptoms, examination findings, laboratory evidence, and the suspected infection source [8]. Professional guidance emphasizes prompt clinical assessment and cautious interpretation of general adult screening approaches in obstetric populations [9]. The WHO Global Maternal Sepsis Study demonstrated substantial variation in maternal infection and its clinical management across health facilities in multiple countries [10]. More recent maternal sepsis guidance continues to emphasize prompt evaluation when infection-related organ dysfunction is suspected [11]. These sources support interpreting physiological findings within the wider obstetric clinical context rather than relying on an isolated abnormal value or a single screening tool [8,11].

The available evidence addresses two distinct clinical purposes. Diagnostic studies assess whether vital signs or screening systems identify infection or sepsis, whereas prognostic studies assess subsequent deterioration among women with suspected or established infection. These applications should not be interpreted as equivalent. Recent reviews of maternal and obstetric early warning systems have examined broad obstetric populations and heterogeneous target outcomes rather than focusing specifically on postpartum sepsis recognition [4,6]. Variation in intended use, thresholds, populations, and reference outcomes therefore leaves uncertainty about the diagnostic performance of these approaches specifically among postpartum women.

The present review distinguishes direct postpartum diagnostic evidence from supportive mixed obstetric, descriptive, and prognostic evidence. It evaluates individual vital signs alongside composite warning systems, focusing on two clinically important patterns: frequent screening alerts among women without confirmed infection and incomplete detection among women with infection or sepsis. Postpartum timing, delivery context, infection source, measurement approach, reference standards, and study design are considered when interpreting apparent performance.

The primary research question was: among women from childbirth through 42 days postpartum, how accurately and reliably do routine maternal vital signs and vital sign-based screening systems identify postpartum maternal infection or sepsis? The primary aim was to evaluate the diagnostic performance and clinical limitations of temperature, heart rate, respiratory rate, blood pressure, oxygen saturation, shock index (the ratio of heart rate to systolic blood pressure), and obstetric warning systems for early postpartum recognition. A secondary objective examined their prognostic value for infection-related deterioration among women with suspected or confirmed infection. Additional objectives were to examine clinically relevant variation in performance and identify priorities for postpartum-specific validation studies.

Review

Methods

Study Design and Reporting Framework

This systematic review was conducted in accordance with the PRISMA 2020 statement [12], and the narrative synthesis was reported with reference to the Synthesis Without Meta-Analysis (SWiM) guideline [13]. The review was not prospectively registered. The eligibility criteria, screening process, data extraction approach, synthesis methods, and risk of bias procedures are reported transparently. The final search was completed on July 13, 2026.

The primary review question was: among women from childbirth through 42 days postpartum, how accurately and reliably do routine maternal vital signs and vital sign-based screening systems identify postpartum maternal infection or sepsis? A secondary question examined whether these measures predicted infection-related deterioration among women with suspected or confirmed infection.

The diagnostic review framework used to define the primary question and eligibility is summarized in Table 1.

Component	Definition
Population	Women from childbirth through 42 completed days postpartum. Mixed antepartum, intrapartum, and postpartum populations were eligible when postpartum data were reported separately or when the study provided clearly relevant supportive evidence applicable to postpartum maternal sepsis.
Index tests	Routine maternal vital signs, including temperature, heart rate, respiratory rate, systolic or diastolic blood pressure, mean arterial pressure, oxygen saturation, and shock index, together with vital sign-based sepsis screening or early warning systems.
Target condition and diagnostic reference standards	Postpartum maternal infection or sepsis assessed against an appropriate clinical, microbiological, or organ dysfunction reference standard. Severe maternal outcome endpoints were treated as prognostic outcomes rather than diagnostic reference standards.
Primary diagnostic outcomes	Sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratios, diagnostic OR, area under the receiver operating characteristic curve, and screen positive rate. True positive, false positive, true negative, and false negative classifications were used only when supported by a prespecified threshold and an appropriate diagnostic reference standard.
Secondary prognostic outcomes	Infection-related deterioration, organ dysfunction, septic shock, intensive care or high dependency admission, mechanical ventilation, major surgery, hysterectomy, prolonged hospital stay, treatment delay, and maternal death.
Eligible study designs	Prospective and retrospective cohort studies, diagnostic accuracy studies, case-control studies, nested case-control studies, cross sectional studies, prediction model development or validation studies, population surveillance studies, and quality improvement or service evaluation studies reporting relevant diagnostic or prognostic data.

TABLE 1: Diagnostic review framework and eligibility definitions.

Diagnostic reference standards and prognostic outcome endpoints were treated as distinct methodological constructs throughout the review.

Eligibility Criteria

Studies were selected according to predefined eligibility criteria relating to the postpartum population, index tests, diagnostic target conditions and reference standards, secondary prognostic outcomes, study design, and availability of relevant data. The criteria were designed to prioritize direct postpartum diagnostic evidence while allowing clinically relevant mixed obstetric and prognostic evidence to be retained as a separate supportive evidence layer. The inclusion and exclusion criteria are summarized in Table 2.

Category	Inclusion criteria	Exclusion criteria
Population and timing	Postpartum women from birth through 42 days after delivery	Studies conducted exclusively during pregnancy or labor without relevant postpartum applicability
Postpartum data	Studies with clearly identifiable postpartum populations or separately reported postpartum subgroup data. Mixed obstetric studies were retained only as supportive evidence when they included postpartum participants and evaluated a directly relevant vital sign-based diagnostic, screening, prediction, or prognostic system.	Studies without relevant postpartum applicability or usable postpartum evidence.
Index tests	Routine maternal vital signs, derived physiological indices, or obstetric sepsis screening or early warning systems, including temperature, heart rate, respiratory rate, blood pressure, oxygen saturation, shock index, SIRS, qSOFA, obstetric modified SOFA, Sepsis in Obstetrics Score, MEOWS, MOEWS, Maternal Early Warning Criteria, MEWS, NEWS, eCART, and other relevant obstetric specific systems.	Studies that did not assess a routine vital sign, derived physiological index, or vital sign-based screening system. Studies focused solely on biomarkers, microbiology, imaging, treatment, antibiotic prophylaxis, prevalence, or infection risk factors without directly extractable diagnostic or prognostic vital sign data.
Outcomes	For diagnostic analyses, postpartum infection or maternal sepsis. For the secondary prognostic component, infection-related deterioration, organ dysfunction, septic shock, intensive care admission, major intervention, or maternal death.	Studies evaluating noninfectious obstetric morbidity without a separate infection, sepsis, or infection-related prognostic outcome.
Reference standards	For diagnostic analyses, an appropriate clinical, microbiological, or organ dysfunction reference standard. Severe maternal outcome endpoints were treated as prognostic outcomes rather than diagnostic reference standards.	Studies lacking an eligible diagnostic reference standard or a clearly defined infection-related prognostic outcome.
Study design and data availability	Eligible study designs reporting extractable data relevant to the index test and diagnostic or prognostic outcome.	Reviews, systematic reviews, meta analyses, editorials, commentaries, letters, case reports, protocols without completed results, qualitative studies, surveys, conference abstracts without sufficient data, animal studies, implementation only studies without eligible diagnostic validation or outcome data, and reports without extractable data relevant to the index test or outcome.
Duplicate or overlapping reports	For multiple publications from the same or substantially overlapping population, the most complete peer reviewed report containing the most relevant postpartum vital sign or outcome data was retained. Companion reports were used only when they provided unique outcomes without duplicating participants or events.	Duplicate or substantially overlapping reports that did not provide unique eligible data.

TABLE 2: Inclusion and exclusion criteria.

eCART, electronic Cardiac Arrest Risk Triage score; MEOWS, Maternal Early Obstetric Warning System; MEWC, Maternal Early Warning Criteria; MEWS, Modified Early Warning Score; MOEWS, Modified Obstetric Early Warning System; NEWS, National Early Warning Score; omSOFA, obstetric modified Sequential Organ Failure Assessment; qSOFA, quick Sequential Organ Failure Assessment; SIRS, systemic inflammatory response syndrome; SOS, Sepsis in Obstetrics Score

Mixed obstetric studies without separate postpartum estimates were considered supportive evidence only and were analyzed separately from studies providing direct postpartum diagnostic data. They were not treated as equivalent sources of postpartum diagnostic accuracy.

The eligibility framework therefore preserved the distinction between evidence used for early diagnostic recognition and evidence used for subsequent severity prediction.

Methodological quality was not used as an eligibility criterion. Risk of bias was assessed after study inclusion and was used only to inform interpretation of the evidence.

Information Sources and Search Strategy

The search strategy combined four concepts: the postpartum population, maternal infection or sepsis,

routine vital signs or obstetric warning systems, and diagnostic or prognostic performance. Database-specific syntax was adapted to each platform. The complete strategies used are presented in Table 3.

Database/source	Search strategy
PubMed/MEDLINE	((("Postpartum Period"[Mesh] OR postpartum[tiab] OR postnatal[tiab] OR puerperal[tiab] OR puerperium[tiab] OR "after childbirth"[tiab] OR "after delivery"[tiab]) AND ("Sepsis"[Mesh] OR "Puerperal Infection"[Mesh] OR "maternal sepsis"[tiab] OR "postpartum sepsis"[tiab] OR "puerperal sepsis"[tiab] OR "postpartum infection"[tiab] OR "maternal infection"[tiab] OR bacteremia[tiab] OR bacteraemia[tiab]) AND ("Vital Signs"[Mesh] OR temperature[tiab] OR fever[tiab] OR hypothermia[tiab] OR tachycardia[tiab] OR "heart rate"[tiab] OR "respiratory rate"[tiab] OR tachypnea[tiab] OR tachypnoea[tiab] OR "blood pressure"[tiab] OR hypotension[tiab] OR "oxygen saturation"[tiab] OR SpO2[tiab] OR "shock index"[tiab] OR SIRS[tiab] OR qSOFA[tiab] OR SOFA[tiab] OR MEOWS[tiab] OR MOEWS[tiab] OR MEWC[tiab] OR MEWS[tiab] OR NEWS[tiab] OR "early warning system"[tiab] OR "early warning score"[tiab] OR "Sepsis in Obstetrics Score"[tiab]) AND (screen*[tiab] OR diagnos*[tiab] OR recogn*[tiab] OR detect*[tiab] OR predict*[tiab] OR sensitivity[tiab] OR specificity[tiab] OR accuracy[tiab] OR "predictive value"[tiab] OR deterioration[tiab] OR "severe maternal outcome"[tiab] OR mortality[tiab]))
Scopus	TITLE-ABS-KEY((postpartum OR postnatal OR puerperal OR puerperium OR "after childbirth" OR "after delivery") AND (sepsis OR "maternal sepsis" OR "postpartum sepsis" OR "puerperal sepsis" OR "postpartum infection" OR "maternal infection" OR bacteremia OR bacteraemia) AND ("vital signs" OR temperature OR fever OR hypothermia OR tachycardia OR "heart rate" OR "respiratory rate" OR tachypnea OR tachypnoea OR "blood pressure" OR hypotension OR "oxygen saturation" OR SpO2 OR "shock index" OR SIRS OR qSOFA OR SOFA OR MEOWS OR MOEWS OR MEWC OR MEWS OR NEWS OR "early warning system" OR "early warning score" OR "Sepsis in Obstetrics Score") AND (screen* OR diagnos* OR recogn* OR detect* OR predict* OR sensitivity OR specificity OR accuracy OR "predictive value" OR deterioration OR "severe maternal outcome" OR mortality))
Web of Science Core Collection	TS=((postpartum OR postnatal OR puerperal OR puerperium OR "after childbirth" OR "after delivery") AND (sepsis OR "maternal sepsis" OR "postpartum sepsis" OR "puerperal sepsis" OR "postpartum infection" OR "maternal infection" OR bacteremia OR bacteraemia) AND ("vital signs" OR temperature OR fever OR hypothermia OR tachycardia OR "heart rate" OR "respiratory rate" OR tachypnea OR tachypnoea OR "blood pressure" OR hypotension OR "oxygen saturation" OR SpO2 OR "shock index" OR SIRS OR qSOFA OR SOFA OR MEOWS OR MOEWS OR MEWC OR MEWS OR NEWS OR "early warning system" OR "early warning score" OR "Sepsis in Obstetrics Score") AND (screen* OR diagnos* OR recogn* OR detect* OR predict* OR sensitivity OR specificity OR accuracy OR "predictive value" OR deterioration OR "severe maternal outcome" OR mortality))
Google Scholar	(postpartum OR postnatal OR puerperal OR puerperium) AND (sepsis OR "maternal sepsis" OR "postpartum sepsis" OR "puerperal sepsis" OR "postpartum infection") AND ("vital signs" OR temperature OR fever OR tachycardia OR "heart rate" OR "respiratory rate" OR hypotension OR "oxygen saturation" OR "shock index" OR SIRS OR qSOFA OR MEOWS OR "early warning score") AND (screening OR diagnosis OR recognition OR detection OR prediction OR sensitivity OR specificity). Exact title and author name searches were used to retrieve potentially eligible reports.

TABLE 3: Search strategies for PubMed/MEDLINE, Scopus, Web of Science Core Collection, and Google Scholar.

Searches covered database inception through July 13, 2026 for PubMed/MEDLINE, Scopus, and Web of Science Core Collection; Google Scholar was searched through July 13, 2026. No publication date restriction was applied, and eligibility was restricted to English-language reports.

MeSH, Medical Subject Headings; tiab, title and abstract; TITLE ABS KEY, title, abstract, and keywords; TS, topic search

Supplementary searching comprised Google Scholar searches, backward reference screening, and forward citation searching. Google Scholar was used to identify potentially eligible studies through combinations of the four review concepts and to retrieve known reports through exact title and author name searches. The first 100 Google Scholar results were screened in the default relevance order. Results became increasingly unrelated to the review question from approximately the 70th result onward, and screening was stopped after the 100th result. This stopping decision was made during screening rather than specified in advance. Backward searching involved screening the reference lists of all included studies and relevant reviews. Forward citation searching sought later studies evaluating the same vital signs, scoring systems, or populations. Reports identified through supplementary searching were subject to the same eligibility criteria and duplicate removal procedures as database records.

Citation details were compared across all sources to identify title variations, incomplete records, duplicate publications, and overlapping cohorts. When the same study was identified from more than one source, only one bibliographic record was retained for screening.

Record Management and Duplicate Removal

All retrieved records were entered into a structured Excel screening spreadsheet (Microsoft Corporation, Redmond, WA, USA). Duplicate records were identified by comparing titles, author names, publication year, DOI, PMID, journal, and publisher records. Minor differences in punctuation, title wording, spelling, or author order were manually reviewed before records were classified as duplicates.

Where the same study appeared in more than one database or supplementary source, only one bibliographic record was retained for screening. When a preliminary report and a later complete peer-reviewed article described the same cohort, the most complete full-text publication was retained.

Study Selection

Study selection was conducted in two stages. First, titles and available abstracts or bibliographic information were screened against the predefined eligibility criteria. Nonoriginal reports, nonhuman studies, and records addressing an ineligible population, index test, outcome, or study design were excluded at this stage.

Second, the full-text of potentially eligible reports was assessed against the predefined eligibility criteria. DOI records, bibliographic database entries, and publisher information were used when needed to verify publication identity and eligibility. Exclusion reasons were assigned according to the categories summarized in Table 2.

Two reviewers independently screened titles, abstracts, and full-text reports. Disagreements were resolved through discussion and consensus. A third reviewer was available to adjudicate unresolved disagreements. A formal inter-reviewer agreement statistic was not calculated.

Verified publications for which a complete report could not be obtained were recorded as reports not retrieved. Records that could not be verified as published reports were excluded during metadata screening and were not counted as full-text reports.

Evidence Layer Classification

Included studies were classified into two evidence layers before synthesis. Layer 1 consisted of direct postpartum diagnostic evidence and included studies restricted to postpartum women or studies reporting separately extractable postpartum diagnostic data. These studies were used to address the primary review question concerning early recognition of postpartum maternal infection or sepsis.

Layer 2 consisted of supportive evidence and included mixed obstetric studies without separate postpartum performance estimates, descriptive studies, and studies evaluating severity or deterioration after infection or sepsis had already been suspected or diagnosed. Layer 2 studies were used to contextualize the interpretation of routine vital signs, early warning tools, and severity prediction but were not treated as equivalent to direct postpartum diagnostic evidence.

Data Extraction

Two reviewers independently extracted data from each included study using a standardized Excel workbook developed for this review. The extraction form was refined during extraction of an initial subset of studies before being applied to the complete evidence set. Extracted data were compared for completeness and consistency, and disagreements were resolved through discussion and consensus.

Extracted study characteristics were grouped into bibliographic and contextual variables, study design variables, and population variables. These included first author, publication year, country, resource setting, clinical setting, study design, prospective or retrospective data collection, recruitment period, number of centers, total sample size, number of postpartum participants, number of infection or sepsis cases, comparison group, where applicable, postpartum timing, mode of delivery, infection source, and evidence layer classification. Resource setting and clinical setting were retained as separate variables because they described different aspects of the study context.

For each individual vital sign or composite tool, the reviewers extracted the index test, tool version, physiological components, measurement method, measurement frequency, threshold or trigger rule, whether the threshold was prespecified, and whether a single, first, worst, peak, minimum, or repeated value was used. Particular attention was given to timing relative to childbirth, postpartum phase, and the timing of sepsis recognition because the clinical meaning of the same physiological abnormality may vary across the postpartum period.

The reference standard was recorded as reported and categorized as microbiologically confirmed infection, clinically adjudicated infection, or maternal sepsis with organ dysfunction. Severe maternal outcome

endpoints, including intensive care admission, septic shock, major intervention, or maternal death, were recorded separately as prognostic outcomes rather than diagnostic reference standards. The reviewers also recorded whether a diagnostic reference standard incorporated any of the vital signs being evaluated.

When an index test was evaluated against an appropriate diagnostic reference standard, extracted performance data included true positive, false positive, true negative, and false negative counts; sensitivity; specificity; positive predictive value; negative predictive value; likelihood ratios; diagnostic OR; area under the receiver operating characteristic curve; screen positive rate; and confidence intervals. When complete 2 by 2 data were available but summary measures were not reported, the review team calculated sensitivity as $TP/(TP + FN)$, specificity as $TN/(TN + FP)$, positive predictive value as $TP/(TP + FP)$, and negative predictive value as $TN/(TN + FN)$, where TP, FN, TN, and FP denote true positives, false negatives, true negatives, and false positives, respectively. These estimates were labelled as reviewer-calculated values to distinguish them from estimates reported in the original studies. No diagnostic classification was assigned when a prespecified threshold or an appropriate diagnostic reference standard was absent.

Descriptive clinical patterns were extracted separately from formal diagnostic performance measures. These included fever or hypothermia without confirmed infection, sepsis without fever, tachycardia among women without sepsis, sepsis without tachycardia, and sepsis occurring without tachypnea, hypotension, or hypoxemia, together with incomplete or selectively recorded physiological measurements. These patterns were not labelled as false positive or false negative findings unless a prespecified threshold and an appropriate diagnostic reference standard permitted that classification.

Secondary prognostic outcomes included infection-related organ dysfunction, septic shock, intensive care or high dependency admission, mechanical ventilation, major surgery, hysterectomy, length of stay, readmission, time from the first abnormal sign to recognition, time to antibiotic treatment, time to source control or transfer, and maternal death.

Study authors were not contacted for missing data. Information that was not reported was recorded as not reported rather than being assumed to represent absence of the characteristic or outcome.

Outcomes

The primary outcomes were formal measures of diagnostic performance and clinical limitations of individual routine vital signs and vital sign-based screening systems for postpartum maternal infection or sepsis. These included sensitivity, specificity, positive and negative predictive values, likelihood ratios, diagnostic OR, area under the receiver operating characteristic curve, and screen positive rates. True positive, false positive, true negative, and false negative classifications were used only when the study evaluated a prespecified threshold against an appropriate diagnostic reference standard. Descriptive patterns, such as abnormal observations without confirmed infection or sepsis occurring without a specified physiological abnormality, were analyzed separately from formal diagnostic accuracy measures.

Secondary outcomes concerned prognosis after suspected or confirmed infection and included infection-related clinical deterioration, organ dysfunction, septic shock, intensive care admission, mechanical ventilation, major surgical intervention, prolonged hospitalization, treatment delay, and maternal death.

Differences in outcome definitions, index test thresholds, postpartum timing, measurement frequency, reference standards, and intended clinical purpose were recorded and considered during synthesis.

Risk of Bias Assessment

Risk of bias was assessed after study inclusion by two reviewers independently, and disagreements were resolved through discussion and consensus. Diagnostic accuracy studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 tool [14]. Risk of bias judgments addressed the internal validity of the reported diagnostic estimates across patient selection, index test, reference standard, and participant flow and timing. Applicability concerns were considered separately and addressed how closely the study population, index test, and reference standard corresponded to the postpartum diagnostic question of this review.

Prediction model development, validation, and prognostic score or marker studies were assessed using the Prediction Model Risk of Bias Assessment Tool [15]. The domains included participant selection, predictor definition and measurement, outcome definition and measurement, and statistical analysis. Particular attention was given to small event numbers, post hoc threshold testing, missing data, overfitting, calibration, internal and external validation, and evaluation of a model or threshold in the same sample in which it was developed.

Cohort, case-control, surveillance, descriptive, and quality improvement studies that were not appropriately classified as diagnostic accuracy or prediction studies underwent a structured qualitative appraisal informed

by the relevant Joanna Briggs Institute critical appraisal criteria [16]. The appraisal considered participant selection, exposure measurement, outcome ascertainment, confounding, completeness of follow-up, and appropriateness of statistical analysis.

Additional concerns affecting internal validity included incorporation of the evaluated vital signs into the sepsis definition, differential verification when only women with abnormal observations underwent complete infection assessment, and selective exclusion of participants with missing score components. Restriction to SIRS positive or intensive care populations was considered in relation to patient selection and the intended use of the test. Mixed obstetric populations and the absence of separately reported postpartum results were considered limitations of applicability to the postpartum review question rather than, by themselves, evidence of high risk of bias. Overlapping cohorts were considered when assessing the independence of evidence contributed to the synthesis.

For QUADAS 2 and PROBAST, domain judgments were recorded as low, high, or unclear risk of bias in accordance with the respective tools, and applicability concerns were recorded separately where relevant. For the JBI-informed qualitative appraisal, study-specific methodological concerns were summarized. Overall categories of low, moderate, or high methodological concern represented review-defined qualitative interpretations of the concerns described for each study. No numerical score or summed threshold was used, and these categories were not formal or validated JBI classifications.

All risk of bias and methodological concern judgments were used only to guide interpretation of the evidence and were not used as exclusion criteria. Individual study judgments are presented by assessment domain. The principal methodological concerns identified across these assessments are summarized in the Results under Risk of Bias.

Data Synthesis

A narrative synthesis was performed in accordance with the SWiM reporting guideline [13]. Meta-analysis was not performed because the included studies differed substantially in postpartum timing, population selection, index tests, thresholds, diagnostic reference standards, prognostic endpoints, intended clinical purpose, sampling methods, study design, and reported performance measures. Comparable 2 by 2 diagnostic data were available for only a limited subset of studies, and pooling clinically different diagnostic and prognostic questions was considered potentially misleading. Reported estimates were compared descriptively within the relevant clinical purpose and evidence layer. No formal between-study statistical tests, indirect comparisons between tools, or pooled estimates were performed. Differences between tools were not interpreted as statistically significant unless such comparisons had been formally tested in the original study.

The synthesis was organized by clinical purpose. Evidence concerning early detection and diagnostic assessment was considered first, beginning with direct postpartum studies and followed by supportive mixed obstetric screening studies. Models intended to identify infection or sepsis were interpreted within the diagnostic component, whereas models intended to predict subsequent deterioration among women with suspected or confirmed infection were interpreted within the prognostic component. Their placement together did not imply an equivalent clinical purpose. Within each component, findings were grouped by individual vital sign or composite tool, including temperature, heart rate, respiratory rate, blood pressure, shock index, oxygen saturation, and obstetric warning systems.

Within each clinical purpose and evidence layer, findings were compared descriptively according to the target condition or outcome, reference standard, index test threshold, and timing of measurement. Sensitivity, specificity, predictive values, and area under the curve were interpreted in relation to the population and outcome evaluated. Estimates addressing infection detection, culture positivity, organ dysfunction, or subsequent severe outcomes were not treated as interchangeable. Formal diagnostic performance measures were also distinguished from descriptive observations of abnormal vital signs without confirmed infection or sepsis without a specified physiological abnormality.

Interpretation considered postpartum phase, mode of delivery, infection source, resource and clinical setting, study design, measurement frequency, missing data, reference standard validity, and risk of bias. Directness to postpartum women, methodological limitations, and consistency across clinically relevant studies informed the strength and scope of the conclusions. Findings from selected or mixed populations were interpreted within those limits. Overlapping reports contributed only once to each outcome.

Assessment of Certainty

A formal GRADE assessment was not performed because the review combined heterogeneous diagnostic accuracy, prediction, prognostic, cohort, surveillance, and descriptive evidence addressing different clinical questions.

Certainty was interpreted narratively according to study design, consistency, directness to postpartum

women, number of studies and events, sample size, reference standard validity, completeness of vital sign measurement, risk of bias, external validation, and heterogeneity in thresholds and outcome definitions. These factors informed the wording of the Discussion and Conclusion.

Results

Study Selection

The search identified 488 records from all search sources. After removal of 33 duplicates, 455 unique records underwent title, abstract, or metadata screening, and 412 were excluded because they did not meet the eligibility criteria. Full-text retrieval was sought for 43 reports. Eleven reports could not be retrieved, leaving 32 full-text reports for eligibility assessment. Thirteen reports were excluded according to the predefined eligibility criteria. Nineteen studies were included in the narrative synthesis [17-35]. Six studies provided direct postpartum diagnostic evidence [17-22], whereas 13 provided supportive mixed population, descriptive, or prognostic evidence [23-35]. This evidence layer distinction was retained throughout the subsequent Results sections.

The study selection process is summarized in Figure 1.

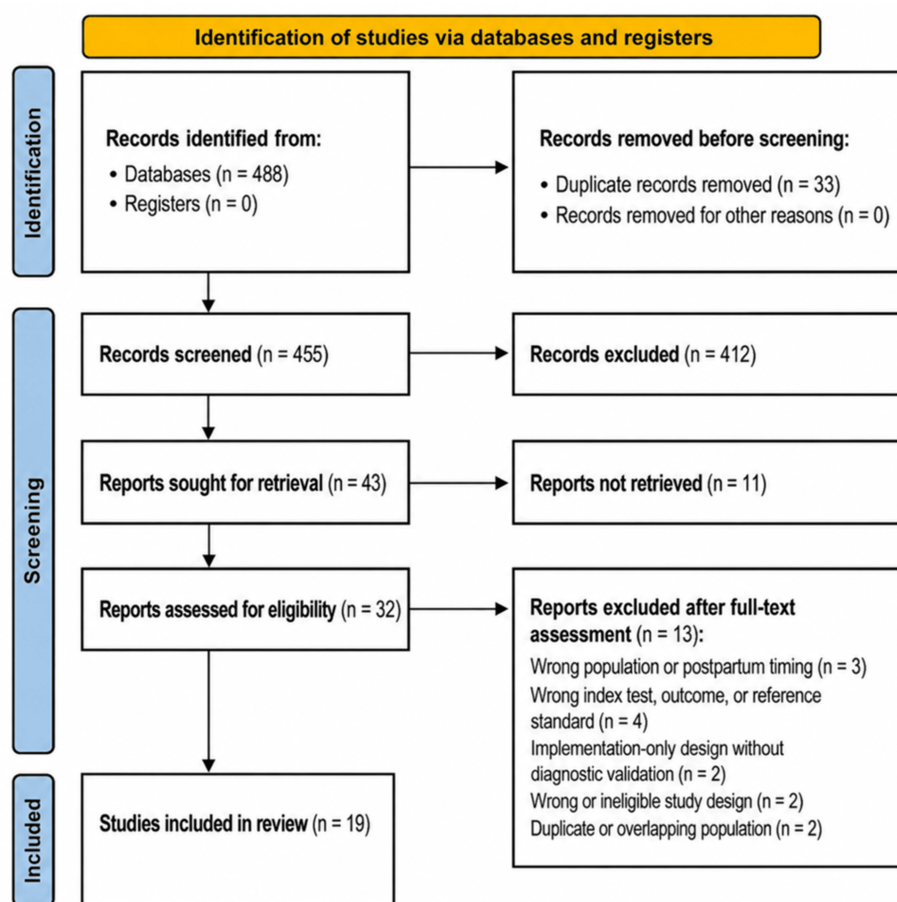


FIGURE 1: PRISMA 2020 flow diagram for study identification, screening, eligibility assessment, and inclusion.

The diagram summarizes the identification, screening, full-text eligibility assessment, exclusion, and final inclusion of 19 studies.

Study Characteristics

The 19 included studies were published between 2018 and 2025 and were conducted across North America, Europe, Africa, and South Asia, with one multinational analysis covering 46 countries [17-35]. Six studies enrolled postpartum-only populations or reported directly extractable postpartum results [17-22]. The remaining studies evaluated mixed obstetric populations, severity after suspected or confirmed infection, or maternal sepsis surveillance [23-35]. Designs included prospective and retrospective cohorts, nested and

matched case-control studies, diagnostic accuracy analyses, prognostic score evaluations, population surveillance, and quality improvement evaluations. Sample sizes varied markedly, from 25 patients in a tertiary intensive care cohort to 19,611 obstetric ward admissions. Postpartum observation windows ranged from the first 24 hours after birth to 42 days postpartum, although several mixed studies did not report postpartum-specific estimates. The evaluated index tests included individual vital signs, shock index, SIRS, qSOFA, SOFA-derived scores, the Sepsis in Obstetrics Score, MEOWS, MOEWS, and MEWC variants, eCART, NICE risk stratification tools, and community clinical algorithms.

The main characteristics of the included studies are summarized in Table 4.

Study	Country/setting	Design	Sample	Postpartum timing	Index test	Reference outcome or endpoint	Evidence role
Shafik et al. (2019) [17]	USA; single hospital	Retrospective consecutive chart review	(n = 638) immediate postpartum women; sepsis cases (n = 8)	0-24 hours	SIRS and individual components	Clinical chart diagnosis of acute postpartum sepsis	Layer 1: direct postpartum
Hannola et al. (2021) [18]	Finland; tertiary hospital	Prospective high-risk cohort	(n = 828) women; postpartum infections (n = 33)	First 24 hours	Obstetric Early Warning System (OEWS)	Composite puerperal infection outcome	Layer 1: direct postpartum
Bakhtawar et al. (2020) [19]	Pakistan; tertiary hospital	Nested case-control/model development	(n = 598) postpartum women; SIRS-defined cases (n = 100); controls (n = 498)	Within 42 days; exact timing unclear	Clinical model including SpO ₂ , symptoms and risk factors	SIRS-defined postpartum sepsis	Layer 1: direct postpartum
Ngonzi et al. (2018) [20]	Uganda; regional referral hospital	Prospective cohort	(n = 4,231) women; fever/hypothermia (n = 205); confirmed infections (n = 84)	In-hospital postpartum	Temperature screening plus structured clinical/microbiological assessment	Endometritis, UTI or bloodstream infection	Layer 1: direct postpartum
Ngunyi et al. (2020) [21]	Cameroon; tertiary referral hospital	Retrospective hospital cohort	(n = 1,520) postpartum records; pyrexia cases (n = 134)	Days 1-10 postpartum	Postpartum temperature/pyrexia	Etiological diagnosis including malaria, UTI and puerperal sepsis	Layer 1: direct postpartum
LeFevre et al. (2021) [22]	Bangladesh and Pakistan; community sites	Nested case-control diagnostic validation	(n = 856) postpartum women across three sites	Community postpartum visits	CHW clinical algorithm: measured/reported fever, pain and discharge	Physician assessment using the same clinical algorithm	Layer 1: direct postpartum
Champagne and Garabedian (2020) [23]	USA; maternity service	Retrospective quality-improvement evaluation	(n = 5,075) births; sepsis episodes (n = 204)	Antepartum, intrapartum, postpartum and readmission	Obstetric-adjusted SIRS pathway	Sepsis defined by positive SIRS screen plus suspected infection	Layer 2: contextual/mixed
Bauer et al. (2019) [24]	USA and Israel; seven academic centers	Multicenter retrospective case-control	Severe sepsis cases (n = 82); controls (n = 328); score analysis (n = 240)	Delivery hospitalization; mixed peripartum	SIRS, qSOFA and modified MEW criteria	Verified severe sepsis with organ failure/hypotension/hypoperfusion	Layer 2: contextual/mixed
Arnolds et al. (2022) [25]	USA; tertiary hospital	Retrospective EHR cohort/external score validation	(n = 19,611) antepartum/postpartum admissions; infections (n = 88)	Ward admission; immediate post-birth period excluded	MEWS, NEWS, eCART, MEOWS, MEWC and MEWT	New infection defined by blood-culture order plus sustained antibiotic treatment	Layer 2: contextual/mixed
Chimwaza et al. (2025) [26]	46 countries; WHO GLOSS cohort	Secondary diagnostic/prognostic validation	(n = 2,560) women with suspected/confirmed infection; severe maternal outcomes (n = 381)	Pregnancy to 42 days postpartum; mixed	21 early warning systems	Infection-related WHO near miss or maternal death	Layer 2: contextual/mixed
Nathan et al. (2019) [27]	South Africa; three tertiary hospitals	Prospective prognostic cohort	(n = 126) maternal sepsis cases; deaths (n = 10); antenatal diagnoses (n = 13)	Predominantly postnatal	First and worst shock index after diagnosis	Death, critical-care admission, major procedure and hysterectomy	Layer 2: prognostic
Aganwal et al. (2019) [28]	India; tertiary obstetric emergency department	Prospective prognostic validation	(n = 100) SIRS-positive women; postpartum (n = 67); severe sepsis (n = 58)	Up to 6 weeks postpartum; mixed	Sepsis in Obstetrics Score (SOS)	Organ failure/severe pregnancy-associated sepsis	Layer 2: prognostic
Reicher et al. (2023) [29]	Israel; tertiary hospital	Retrospective postpartum infection cohort	(n = 57) culture-confirmed postpartum GAS infections; ICU/STSS (n = 12)	Early and late postpartum	Fever, tachycardia, CRP and treatment delay	STSS and ICU admission	Layer 2: prognostic
Kuhr et al. (2025) [30]	Sierra Leone; three hospitals	Prospective prognostic cohort	(n = 855) maternal sepsis cases; deaths (n = 43)	Antepartum/postpartum mixed	First and worst shock index	Death, major procedure, severe anemia and transfusion	Layer 2: prognostic
Dohare et al. (2025) [31]	India; tertiary hospital	Prospective observational score study	(n = 225) SIRS-positive women; puerperal (n = 117)	Antenatal and within 42 days postpartum	Sepsis in Obstetrics Score	Maternal morbidity, transfer, MODS and death	Layer 2: prognostic

Abutheraa et al. (2021) [32]	Scotland; three maternity hospitals	Retrospective diagnostic cohort	Suspected sepsis (n = 89); complete cases (n = 75); positive cultures (n = 29)	Mixed maternity population	mSIRS, qSOFA and MEWS	Positive microbiological culture	Layer 2: contextual/mixed
Karatas et al. (2025) [33]	Turkey; tertiary ICU	Retrospective prognostic cohort	(n = 25) ICU patients with maternal sepsis; after delivery (n = 10); septic shock (n = 3)	Pregnancy/puerperium mixed	SOS and obstetric-modified SOFA	ICU stay >72 hours and septic shock	Layer 2: prognostic
Omaghi et al. (2023) [34]	Italy; six-region population surveillance	Prospective population cohort + nested case-control	Peripartum sepsis (n = 219); postpartum cases (n = 150); controls (n = 297)	Up to 42 days postpartum	Routine signs/symptoms and organ-failure criteria	Maternal sepsis; serious maternal adverse events	Layer 2: contextual/descriptive
Kvalvik et al. (2024) [35]	Norway; tertiary hospital	Retrospective cohort + code-accuracy analysis	(n = 5,182) deliveries; infections (n = 402); maternal sepsis (n = 18); postpartum sepsis (n = 13)	Up to 42 days postpartum	Routine vital signs and ICD-10 O85 code	Expert-adjudicated maternal/puerperal sepsis	Layer 2: contextual/descriptive

TABLE 4: Characteristics of the included studies.

CHW, community health worker; eCART, electronic Cardiac Arrest Risk Triage; EHR, electronic health record; GAS, group A Streptococcus; GLOSS, Global Maternal Sepsis Study; MEW/MEWS, Modified Early Warning/Modified Early Warning Score; MEWC, Maternal Early Warning Criteria; MEOWS/MOEWS, Maternal/Modified Obstetric Early Warning System; NEWS, National Early Warning Score; OEWS, Obstetric Early Warning System; qSOFA, quick Sequential Organ Failure Assessment; SIRS, systemic inflammatory response syndrome; SOS, Sepsis in Obstetrics Score; SpO₂, peripheral oxygen saturation; STSS, streptococcal toxic shock syndrome

Key Findings

Direct postpartum diagnostic and screening evidence: The six direct postpartum studies demonstrated substantial variation in the performance of routine vital signs and screening tools [17-22]. In the immediate postpartum cohort reported by Shafik et al. (2019) [17], 460 of 638 women (72.1%) met SIRS criteria, although only eight women (1.25%) had sepsis. All eight sepsis cases screened positive, but 452 women without sepsis also met SIRS criteria, corresponding to reviewer-calculated sensitivity of 100%, specificity of 28.3%, positive predictive value of 1.7%, and negative predictive value of 100%. Abnormal temperature was the only individual SIRS component independently associated with sepsis. In Hannola et al. (2021) [18], an Obstetric Early Warning System red trigger had sensitivity of 51.5%, specificity of 83.2%, positive predictive value of 14.6%, and negative predictive value of 96.8% for postpartum infection; yellow triggers had sensitivity of 12.1%. SIRS produced frequent alerts in the Shafik cohort, whereas red warning triggers identified approximately half of postpartum infections in the Hannola cohort.

The postpartum model reported by Bakhtawar et al. (2020) [19] had an area under the curve of 0.84, and SpO₂ below 93% was associated with the SIRS-defined sepsis outcome. At the high sensitivity cutoff, the model had sensitivity of 82.0%, specificity of 64.8%, positive predictive value of 31.9%, and negative predictive value of 94.7%. In Ngonzi et al. (2018) [20], 205 of 4,176 women with temperature data developed fever or hypothermia; among the 174 women who underwent complete assessment, 84 (48%) had a confirmed infection, and 90 (52%) had no identified infectious source. In Ngunyi et al. (2020) [21], postpartum pyrexia occurred in 134 of 1,520 women, and puerperal sepsis accounted for 17.9% of pyrexia cases. The two cohorts reported different causes of abnormal postpartum temperature. The community algorithm evaluated by LeFevre et al. (2021) [22] had adjusted sensitivity of 82% and specificity of 90%, with site-specific performance varying across Bangladesh and Pakistan.

The direct postpartum diagnostic and screening findings are summarized in Table 5.

Study	Key quantitative findings	Observed performance and methodological considerations
Shafik et al. (2019) [17]	SIRS positive in 460/638 (72.1%); sepsis in 8/638 (1.25%). All 8 cases met SIRS, but 452/630 women without sepsis also met SIRS. Reviewer-calculated sensitivity 100%, specificity 28.3%, PPV 1.7%, NPV 100%. Abnormal temperature was the only SIRS component independently associated with sepsis (OR 25.90, 95% CI 3.17-211.52).	High screen-positive rate and low specificity in the immediate postpartum period.
Hannola et al. (2021) [18]	For 33 infections, any red OEWS trigger had sensitivity 51.5%, specificity 83.2%, PPV 14.6% and NPV 96.8%. Yellow triggers had sensitivity 12.1% and PPV 2.7%. A red heart-rate trigger was independently associated with infection (adjusted OR 3.25, 95% CI 1.03-10.29).	Red triggers identified approximately half of infections; positive predictive value was low.
Bakhtawar et al. (2020) [19]	The multivariable model had AUC 0.84 (95% CI 0.80-0.89). SpO ₂ <93% was strongly associated with SIRS-defined sepsis (adjusted OR 13.0, 95% CI 4.80-37.10). At the high-sensitivity cutoff, sensitivity was 82.0%, specificity 64.8%, PPV 31.9% and NPV 94.7%.	Model performance was reported in a direct postpartum population; the SIRS-based outcome definition and treatment of missing SpO ₂ values increased risk of bias.
Ngonzi et al. (2018) [20]	Temperature was recorded in 4,176/4,231 (99%); 205 women (5%) developed fever >38 °C or hypothermia <36 °C. Among 174 fully evaluated abnormal-temperature women, 84 (48%) had a confirmed infection and 90 (52%) had no source. Overall infection incidence was 84/4,231 (2%).	Abnormal temperature had limited rule-in value; equivalent reference-standard testing was not applied to normothermic women.
Ngunyi et al. (2020) [21]	Postpartum pyrexia occurred in 134/1,520 women (8.82%). The reported causes were malaria 46.7%, UTI 18.7%, puerperal sepsis 17.9%, and pneumonia 8.7%; multiple causes were possible. Most pyrexia occurred in the first 3 days postpartum.	Pyrexia had several infectious etiologies, and puerperal sepsis accounted for a minority of cases.
LeFevre et al. (2021) [22]	Across three community sites, adjusted sensitivity of the CHW algorithm was 82% and specificity 90% (site sensitivity 78%-95%; specificity 85%-95%). Agreement was highest for measured fever (κ = 0.84-0.97) and lowest for lower abdominal pain.	Performance varied by site; the physician reference assessment used substantially overlapping clinical criteria.

TABLE 5: Direct postpartum diagnostic and screening evidence.

AUC, area under the receiver operating characteristic curve; CHW, community health worker; NPV, negative predictive value; OEWS, Obstetric Early Warning System; PPV, positive predictive value; SIRS, systemic inflammatory response syndrome; SpO₂, peripheral oxygen saturation

Mixed obstetric screening and early warning tools: Five mixed population studies evaluated obstetric screening systems or early warning tools [23-26,32]. In Champagne and Garabedian (2020) [23], 204 sepsis episodes were identified among 5,075 births; 35 episodes occurred postpartum, 10 occurred during postpartum readmission, and 20 were identified from immediately before delivery through three hours after birth. Fourteen combinations of SIRS abnormalities were observed, and postpartum-specific diagnostic performance was not reported. Bauer et al. (2019) [24] reported sensitivity and specificity of 93% and 63% for SIRS, 50% and 95% for qSOFA, and 82% and 87% for modified maternal early warning criteria.

The remaining mixed population studies addressed different clinical outcomes, and their performance estimates were therefore not treated as directly comparable. Arnolds et al. (2022) [25] evaluated detection of new infection and reported an area under the curve of 0.79 for heart rate alone, compared with 0.77 for eCART and 0.71 for both MEWS and NEWS. The WHO GLOSS analysis evaluated prediction of severe maternal outcomes among women with suspected or confirmed infection; the NICE postpartum or post-abortion tool had sensitivity of 97.6% and specificity of 11.8%, modified shock index had sensitivity of 95.7% and specificity of 5.9%, omqSOFA had sensitivity of 38.0% and specificity of 90.4%, and qSOFA had sensitivity of 35.6% and specificity of 83.1% [26]. Abutheraa et al. (2021) [32] used positive microbiological culture as the reference outcome and reported areas under the curve of 0.524 for mSIRS, 0.497 for MEWS, and 0.301 for qSOFA; qSOFA identified no culture-positive case.

The performance of screening tools in mixed obstetric populations is summarized in Table 6.

Study	Main quantitative findings	Observed performance and methodological considerations
Champagne and Garabedian (2020) [23]	204 sepsis episodes among 5,075 births; 35 occurred postpartum, 10 during postpartum readmission and 20 from immediately before delivery to three hours after birth. Fourteen different SIRS combinations occurred and no combination uniquely indicated severity. Postpartum-specific test performance was not reported.	Postpartum-specific diagnostic performance was not reported, and multiple SIRS combinations were observed.
Bauer et al. (2019) [24]	SIRS: sensitivity 93%, specificity 63%; qSOFA: 50% and 95%; modified MEW: 82% and 87%. Analysis included 51 cases and 189 controls after substantial exclusions.	SIRS had higher sensitivity and lower specificity; qSOFA had lower sensitivity and higher specificity; modified MEW criteria showed intermediate performance.
Arnolds et al. (2022) [25]	For new infection, eCART AUC was 0.77 (95% CI 0.75-0.78), MEWS 0.71 and NEWS 0.71. Heart rate alone had AUC 0.79 (95% CI 0.77-0.80), exceeding all complete scores.	Heart rate alone had an AUC similar to or higher than complete scores; immediate post-birth observations were excluded.
Chimwaza et al. (2025) [26]	NICE postpartum/post-abortion tool: sensitivity 97.6%, specificity 11.8%, DOR 5.56. mSIRS: 89.2% and 27.4%. Modified shock index: 95.7% and 5.9%. omqSOFA: 38.0% and 90.4%; qSOFA: 35.6% and 83.1%.	High-sensitivity tools had low specificity, whereas qSOFA-derived tools had higher specificity and lower sensitivity.
Abutheraa et al. (2021) [32]	Against positive culture: mSIRS sensitivity 65.5%, specificity 39.1%, PPV 40.4%, NPV 64.3%, AUC 0.524; MEWS 69.0%, 30.4%, 38.5%, 60.9%, AUC 0.497; qSOFA sensitivity 0%, specificity 95.7%, AUC 0.301.	Discrimination against positive culture was poor; qSOFA identified no culture-positive cases.

TABLE 6: Screening and early-warning tools in mixed obstetric populations.

AUC, area under the receiver operating characteristic curve; DOR, diagnostic OR; eCART, electronic Cardiac Arrest Risk Triage; MEW/MEWS, Modified Early Warning/Modified Early Warning Score; mSIRS, maternity systemic inflammatory response syndrome; NICE, National Institute for Health and Care Excellence; NPV, negative predictive value; omqSOFA, obstetric-modified quick Sequential Organ Failure Assessment; PPV, positive predictive value; qSOFA, quick Sequential Organ Failure Assessment; SIRS, systemic inflammatory response syndrome

Sepsis identification model and prognostic evidence: The seven studies summarized in Table 7 included one model for identifying postpartum sepsis and six studies evaluating deterioration after suspected or confirmed infection [19,27-31,33]. In Nathan et al. (2019) [27], using a first recorded shock index of at least 0.9 as the positive threshold for maternal death yielded sensitivity of 80.0%, specificity of 50.4%, and positive predictive value of 12.3%. A value below 0.9 had a negative predictive value of 96.7%. Kuhrt et al. (2025) [30] evaluated shock index in a larger maternal sepsis cohort and also reported performance using the worst recorded value during the clinical course. Shock index had an area under the curve of 0.76 for death. Using a worst recorded shock index of at least 0.9 as the positive threshold for maternal death yielded sensitivity of 90.7% and specificity of 19.8%. At a threshold of at least 1.7, sensitivity was 39.5%, and specificity was 95.9%; 17 of 50 women in the latter category died. These estimates reflect different measurement time points and should therefore be interpreted as distinct prognostic assessments rather than interchangeable thresholds.

Study	Population/outcome	Predictor	Main performance	Methodological considerations
Bakhtawar et al. (2020) [19]	Postpartum SIRS-defined sepsis	SpO ₂ <93% and a multivariable clinical model	AUC 0.84; sensitivity 82.0% and NPV 94.7% at the high-sensitivity cutoff.	Direct postpartum model; outcome incorporated vital sign criteria and external validation was not reported.
Nathan et al. (2019) [27]	126 maternal sepsis cases; 10 deaths	First and worst shock index after diagnosis	For maternal death, using first SI ≥0.9 as the positive threshold: sensitivity 80.0%, specificity 50.4%, PPV 12.3%, NPV 96.7%. Two of six women with worst SI ≥1.7 died.	Assessed severity after diagnosis rather than initial sepsis recognition.
Agarwal et al. (2019) [28]	100 SIRS-positive women; 67 postpartum; 58 severe sepsis	SOS ≥6	Sensitivity 68.9%, specificity 80.9%, PPV 83.3%, NPV 65.3%, AUC 0.810 for organ failure/severe sepsis.	Postpartum-specific performance was not reported separately.
Reicher et al. (2023) [29]	57 culture-confirmed postpartum GAS infections; 12 ICU/STSS	Tachycardia, CRP and treatment delay	Tachycardia occurred in 50.0% of ICU patients versus 17.7% of non-ICU patients (p = 0.02). Delayed antibiotic treatment (>24 hours) was also more frequent among severe cases.	Small organism-specific cohort; predictor analyses were mainly univariable.
Kuhr et al. (2025) [30]	855 maternal sepsis cases; 43 deaths	Worst SI ≥0.9 and ≥1.7	For death, SI AUC 0.76. Worst SI ≥0.9: sensitivity 90.7%, specificity 19.8%, PPV 5.7%, NPV 97.6%. Worst SI ≥1.7: sensitivity 39.5%, specificity 95.9%, PPV 34.0%; 17/50 women died.	Performance varied by threshold; postpartum-specific estimates were not reported separately.
Dohare et al. (2025) [31]	225 SIRS-positive women; 117 puerperal	SOS >6	139/225 had scores >6. Higher scores were associated with transfer (6.8% vs 1.7%), mortality (1.8% vs 0.6%), and adverse fetal outcomes.	Postpartum-specific validation was not reported, and some denominators were internally inconsistent.
Karatas et al. (2025) [33]	25 ICU maternal sepsis patients; 3 septic shock	SOS and omSOFA	SOS predicted ICU stay >72 h: AUC 0.865, sensitivity 75%, specificity 100%. omSOFA predicted septic shock: AUC 0.985, sensitivity 100%, specificity 95%.	Very small ICU cohort with three septic shock events and no external validation.

TABLE 7: Postpartum sepsis identification model and prognostic evidence.

AUC, area under the receiver operating characteristic curve; GAS, group A Streptococcus; MODS, multiple organ dysfunction syndrome; NPV, negative predictive value; omSOFA, obstetric-modified Sequential Organ Failure Assessment; PPV, positive predictive value; SI, shock index; SIRS, systemic inflammatory response syndrome; SOS, Sepsis in Obstetrics Score; SpO₂, peripheral oxygen saturation; STSS, streptococcal toxic shock syndrome

The Sepsis in Obstetrics Score showed variable prognostic performance. Agarwal et al. (2019) [28] reported sensitivity of 68.9%, specificity of 80.9%, positive predictive value of 83.3%, negative predictive value of 65.3%, and an area under the curve of 0.810 for severe sepsis or organ failure at a cutoff of at least six. Dohare et al. (2025) [31] reported more transfers, deaths, and adverse fetal outcomes among women with scores above six, but postpartum-specific performance was not presented. In the ICU cohort reported by Karatas et al. (2025) [33], the Sepsis in Obstetrics Score predicted ICU stay longer than 72 hours with an area under the curve of 0.865, while obstetric-modified SOFA predicted septic shock with an area under the curve of 0.985. In culture-confirmed postpartum group A Streptococcus infection, tachycardia occurred in 50.0% of ICU patients and 17.7% of non-ICU patients [29].

The postpartum sepsis identification model and prognostic findings are summarized in Table 7.

Cross-Study Findings by Vital Sign

Temperature was the most frequently evaluated individual vital sign [17,20-22,34,35]. In the Ugandan cohort, 52% of fully assessed women with abnormal temperature had no identified infectious source, while in the Cameroonian cohort puerperal sepsis accounted for 17.9% of postpartum pyrexia cases [20,21]. Conversely, 26.5% of women with peripartum sepsis in the Italian surveillance cohort did not have a temperature above 38 °C, and two of 18 maternal sepsis cases in the Norwegian cohort remained below 38 °C [34,35].

Heart rate showed a different and more variable pattern. Tachycardia contributed to immediate postpartum SIRS positivity, was independently associated with postpartum infection in Hannola et al. (2021) [18], had an area under the curve of 0.79 for new infection in Arnolds et al. (2022) [25], and was more frequent among severe group A *Streptococcus* cases [29]. However, only 45.7% of Italian sepsis cases had heart rate above 100 beats per minute [34].

Respiratory rate was included in SIRS, qSOFA, and several obstetric early warning systems, but evidence for its independent diagnostic contribution was limited. Scores containing respiratory rate thresholds showed low sensitivity in several mixed obstetric studies, and respiratory rate was missing in 22.2% of sepsis cases in the Norwegian cohort [24-26,32,35].

Blood pressure was difficult to interpret because the included populations contained competing obstetric conditions that can alter blood pressure, including hemorrhage, anesthesia exposure, fluid shifts, and hypertensive disorders. Hypotension was not documented in 32.4% of cases in the Italian surveillance cohort, and blood pressure findings varied across the other included settings [18,24,27,30,34,35].

Shock index showed a clearer threshold trade-off, primarily as a prognostic measure rather than a stand-alone diagnostic test. Lower shock index or modified shock index thresholds produced higher sensitivity with low specificity, whereas higher thresholds were more specific but less sensitive [26,27,30]. Interpretation also depended on whether the first or worst recorded value was used.

Oxygen saturation was associated with SIRS-defined or severe sepsis outcomes in selected cohorts, including the Bakhtawar and Agarwal studies [19,28]. However, SpO₂ was selectively or incompletely measured in several studies, including reports in which it was not routinely available [24,35].

The cross-study findings for individual vital signs are summarized in Table 8.

Vital sign	Studies	Findings affecting specificity or interpretation	Findings affecting sensitivity or case detection	Observed cross-study pattern
Temperature	[17,20-22,34,35]	Fever or hypothermia was frequent without confirmed infection: 52% of fully assessed women with abnormal temperature in Uganda had no identified source; pyrexia in Cameroon was more often attributed to malaria or UTI than to puerperal sepsis.	Sepsis could occur without fever: 26.5% of Italian cases lacked temperature above 38 °C; 2 of 18 Norwegian sepsis cases remained below 38 °C.	Temperature was associated with sepsis in some studies, but abnormal temperature also occurred without confirmed infection and sepsis could occur without fever.
Heart rate	[17,18,25,29,34,35]	Postpartum tachycardia contributed to low SIRS specificity and did not independently identify sepsis in Shafik et al.	Tachycardia was not universal among women with infection or sepsis: 45.7% of Italian sepsis cases had heart rate above 100 beats per minute, and 26.8% of GAS cases had tachycardia.	Heart rate showed variable diagnostic performance and also showed associations with severity in selected cohorts.
Respiratory rate	[24-26,32,35]	High-sensitivity tools containing respiratory rate thresholds had low specificity in mixed obstetric cohorts.	Respiratory rate was frequently missing or poorly recorded; 22.2% of values were missing in Kvalvik et al. (2024) [35]. qSOFA-derived tools had low sensitivity despite respiratory rate thresholds.	Respiratory rate evidence was limited by incomplete measurement and low sensitivity of qSOFA-derived tools.
Blood pressure	[18,24,27,30,34,35]	Competing obstetric conditions could alter blood pressure, including hemorrhage, anesthesia exposure, fluid shifts, and hypertensive disorders.	Hypotension was not universal among women with sepsis; 32.4% of cases in the Italian cohort had no documented hypotension.	Isolated blood pressure thresholds showed limited consistency in mixed obstetric settings; shock index provided additional prognostic information.
Shock index	[26,27,30]	Low threshold SI or MSI strategies had very low specificity, including MSI specificity of 5.9% in GLOSS and specificity of 19.8% for worst SI ≥0.9 in Sierra Leone.	Higher shock index thresholds had lower sensitivity; worst SI of at least 1.7 had sensitivity of 39.5% for death in Sierra Leone.	Lower thresholds increased sensitivity and reduced specificity, whereas higher thresholds increased specificity and reduced sensitivity; performance also depended on measurement timing.
Oxygen saturation	[19,24,26,28,35]	SpO ₂ was missing or selectively measured in several cohorts; Bauer et al. (2019) [24] reported that it was not routinely available.	Hypoxemia was not reported as a universal finding among women with sepsis, and the proportion of sepsis cases occurring without hypoxemia was not consistently available.	Low SpO ₂ was associated with severe outcomes in selected cohorts, but selective or incomplete measurement increased risk of bias.

TABLE 8: Cross-study findings by individual vital sign.

AOR, adjusted OR; AUC, area under the receiver operating characteristic curve; BP, blood pressure; GAS, group A Streptococcus; HR, heart rate; MSI, modified shock index; qSOFA, quick Sequential Organ Failure Assessment; SI, shock index; SIRS, systemic inflammatory response syndrome; SpO₂, peripheral oxygen saturation

Risk of Bias

Risk of bias was assessed after study inclusion using design-appropriate tools. QUADAS 2 was applied to diagnostic accuracy studies, PROBAST to prediction or prognostic studies, and JBI-informed appraisal to cohort, surveillance, and quality improvement studies [17-35]. The dominant finding was substantial

methodological concern across the evidence base. All five QUADAS 2 studies had high overall risk of bias, primarily because of non-independent reference standards, enriched or case-control sampling, incomplete score data, or incorporation of index test components into the sepsis definition. All nine PROBAST-assessed studies had high overall risk of bias, with concerns including small event counts, evaluation of predictors or thresholds in the same sample, and limitations in statistical analysis, calibration, or validation. Mixed antepartum and postpartum populations additionally limited applicability to the postpartum review question. In the review-defined qualitative summary of the JBI appraisals, one study had low overall concern, two had moderate concern, and two had high concern. These categories were used for narrative interpretation and were not treated as validated JBI risk of bias scores. No study was excluded because of its risk of bias judgment.

The QUADAS-2 assessment is presented in Figure 2.

Study	Patient selection	Index test	Reference standard	Flow/timing	Overall risk	Applicability	Detailed rationale
Shafik S et al., 2019 [17]	●	●	●	●	●	●	Consecutive postpartum cohort, but the chart-based sepsis diagnosis was not independently standardized and may have been influenced by SIRS findings; only eight sepsis events were observed.
Hannola K et al., 2021 [18]	●	●	●	●	●	●	Prospective measurements were complete, but the infection outcome included temperature and treatment decisions that could have been influenced by the early-warning findings; the cohort was restricted to high-risk women.
LeFevre AE et al., 2021 [22]	●	●	●	●	●	●	Nested case-control sampling, partial verification, and physician use of essentially the same symptoms and signs as the community algorithm created selection and incorporation concerns.
Bauer ME et al., 2019 [24]	●	●	●	●	●	●	Severe cases were compared with healthy controls; missing score components and substantial exclusions affected the index-test analysis, and postpartum performance was not reported separately.
Abutheraa N et al., 2021 [32]	●	●	●	●	●	●	Only treated or suspected cases were sampled, positive culture was an insensitive reference standard, 14 women had incomplete score data, and the postpartum subgroup was not separated.

FIGURE 2: QUADAS-2 risk-of-bias assessment of diagnostic accuracy studies.

Green = low risk of bias or low applicability concern; red = high risk of bias or high applicability concern; gray = unclear risk of bias or unclear applicability concern. QUADAS-2 domains include patient selection, index test, reference standard, and flow and timing. Applicability concerns were assessed separately from risk-of-bias judgments.

The PROBAST assessment is presented in Figure 3.

Study	Participants	Predictors	Outcome	Analysis	Overall risk	Applicability	Detailed rationale
Bakhtawar S et al., 2020 [19]	H	H	H	H	H	L	Case-control model development, SIRS incorporation into the outcome, missing SpO2 values coded as normal, and absence of external validation.
Arnolds DE et al., 2022 [25]	L	L	H	H	H	H	The infection outcome was an electronic surrogate based on culture ordering and sustained antibiotics; repeated observations, calibration, and immediate postpartum applicability were incompletely addressed.
Chimwasa Y et al., 2025 [26]	L	H	L	H	H	U	Extensive missing predictor data and evaluation of numerous scores without independent validation or calibration limited model-performance inference in the mixed obstetric cohort.
Nathan HL et al., 2019 [27]	U	L	L	H	H	U	Only 10 deaths and six women in the highest worst-shock-index category produced imprecise estimates; analyses were largely unadjusted and the population was predominantly, but not exclusively, postnatal.
Agarwal R et al., 2019 [28]	H	L	L	H	H	U	The convenience cohort was restricted to SIRS-positive women, had limited size, and evaluated the SOS cutoff in the same sample without separate postpartum validation.
Reicher I et al., 2023 [29]	H	U	L	H	H	H	The organism-specific cohort contained 57 cases and 12 severe outcomes; predictor analyses were mainly univariable and timing of measurements varied.
Kuhrt K et al., 2025 [30]	U	L	L	H	H	H	Worst shock index was measured during the clinical course and may represent established deterioration; postpartum-specific timing and performance were not reported separately.
Dohare R et al., 2025 [31]	H	H	H	H	H	H	Outcome definitions and some denominators were internally inconsistent, calibration was not reported, and antenatal and postpartum performance was not separated.
Karatas E et al., 2025 [33]	U	L	U	H	H	H	Retrospective ICU cohort of 25 women with three septic-shock events. Thresholds were selected by Youden index and assessed in the same sample without validation. Predictor-outcome overlap and blinding were unclear; the selected mixed population limited postpartum applicability.

FIGURE 3: PROBAST risk-of-bias assessment of prediction and prognostic score studies.

Green = low risk of bias or low applicability concern; red = high risk of bias or high applicability concern; gray = unclear risk of bias or unclear applicability concern. PROBAST domains include participants, predictors, outcome, and analysis. Applicability concerns were assessed separately from the overall risk-of-bias judgment.

The JBI-informed appraisal is presented in Figure 4.

Study	Sampling/selection	Exposure measurement	Outcome ascertainment	Confounding/follow-up	Overall concern	Detailed rationale
Ngonzi J et al., 2018 [20]	L	H	H	L	H	Temperature was recorded prospectively, but only febrile or hypothermic women underwent complete microbiological and clinical verification, producing differential verification bias.
Ngunyi YL et al., 2020 [21]	H	H	H	H	H	Retrospective referral-hospital records, non-standardized diagnostic investigation, and possible multiple aetiologies limited outcome classification.
Champagne HA and Garabedian MJ, 2020 [23]	H	L	H	H	H	Sepsis episodes were operationally identified through the SIRS pathway itself; there was no non-sepsis comparator and no postpartum-specific accuracy analysis.
Ornaghi S et al., 2023 [34]	L	L	L	L	L	Prospective population surveillance used expert criteria and a nested postpartum case-control analysis; some regions were excluded for underreporting and vital-sign findings were not reported separately for postpartum cases.
Kvalvik SA et al., 2024 [35]	H	L	L	H	H	Expert chart adjudication strengthened case classification, but ICD-based case finding could miss uncoded infections and only 18 sepsis cases were available.

FIGURE 4: JBI-informed critical appraisal of cohort, surveillance, and quality-improvement studies.

Green = low methodological concern; orange = moderate methodological concern; red = high methodological concern; gray = unclear methodological concern. The appraisal considered participant selection, exposure measurement, outcome ascertainment, confounding, follow-up, and appropriateness of statistical analysis. Overall methodological concern categories were review-defined qualitative summaries and were not validated JBI scores.

Discussion

This systematic review synthesized 19 studies evaluating routine vital signs, obstetric early warning systems, sepsis screening tools, and derived physiological indices in relation to postpartum maternal infection, sepsis, or infection-related deterioration [17-35]. Among studies addressing early detection, the main challenge was balancing screening alert burden against incomplete case detection. The evidence did not establish a consistently reliable isolated vital sign, threshold, or score for confirming or excluding postpartum maternal sepsis. Interpretation depended on postpartum timing, measurement quality, clinical context, and whether the intended purpose was diagnosis or prognosis.

The six studies providing direct postpartum evidence reduced reliance on extrapolation from mixed obstetric populations [17-22]. Shafik et al. demonstrated the limited specificity of SIRS immediately after childbirth [17], whereas Hannola et al. demonstrated incomplete detection of postpartum infection using red warning triggers [18]. These findings illustrate different screening limitations rather than establish the superiority of one approach. Their clinical implications depend on the consequences of additional assessments, missed infections, and delayed escalation in the setting where a tool is used.

Temperature findings also require interpretation according to the population and outcome studied. Shafik et al. evaluated immediate postpartum sepsis [17], the Ugandan and Cameroonian cohorts examined abnormal temperature within broader postpartum infection patterns [20,21], and the Italian and Norwegian studies described physiological findings among women with maternal sepsis [34,35]. Collectively, these studies indicate that abnormal temperature does not necessarily establish sepsis and that the absence of fever does not exclude it. They do not provide a single comparable estimate of temperature accuracy across postpartum settings. Interpretation should also consider postpartum phase, local infection patterns, delivery context, and prior treatment.

Other vital signs showed similarly variable associations. Heart rate contributed to screening alerts but was also associated with infection or greater severity in selected populations [17,18,25,29]. Evidence concerning respiratory rate was limited by incomplete recording and by studies evaluating composite scores rather than its independent contribution [24-26,32,35]. Blood pressure findings required consideration of both infection-related deterioration and competing obstetric conditions, while hypotension was not documented in all sepsis cases [18,24,27,30,34,35]. Oxygen saturation findings were constrained by selective measurement, missing observations, and overlap between predictors and outcome definitions [19,28,35]. These limitations prevent confident attribution of diagnostic performance to individual components of a composite score.

Composite screening results likewise depended on the target condition. In the cited mixed obstetric studies, SIRS or maternity SIRS had higher sensitivity and lower specificity than qSOFA, but one study evaluated sepsis and the other evaluated positive microbiological culture [24,32]. Obstetric modified early warning criteria showed intermediate sensitivity and specificity in the multicenter sepsis study [24]. These findings describe performance within the respective study populations and reference standards. They should not be interpreted as evidence of a universally preferable score or threshold.

The shock index studies primarily informed prognosis after infection or sepsis had been suspected or established. Nathan et al. [27] and Kuhrt et al. [30] evaluated adverse outcomes using measurements taken at different points in the clinical course. A first recorded value may support assessment at presentation, whereas the worst recorded value incorporates information from subsequent deterioration and may be less useful for early prediction. The WHO GLOSS analysis similarly evaluated warning systems against severe maternal outcomes rather than initial infection detection [26]. These distinctions limit the use of prognostic findings to justify claims about early postpartum diagnostic accuracy.

Reference standards were another major source of uncertainty. Positive culture, clinically adjudicated infection, infection-related organ dysfunction, intensive care admission, shock, and death represent different target conditions or outcomes. Culture positivity alone does not establish sepsis, and a negative culture does not exclude clinically important infection. Incorporation of physiological predictors into outcome definitions, incomplete verification, selective measurement, and limited external validation further weakened interpretation. Narrative synthesis allowed these differences to remain visible, although it could not resolve the underlying limitations or establish comparative effectiveness between tools.

The clinical implication is that vital signs should contribute to a broader assessment rather than determine diagnosis independently. Complete observations, reassessment over time, and comparison with the woman's baseline may help clinicians recognize evolving abnormalities. Interpretation should incorporate symptoms, examination findings, delivery context, suspected infection source, competing obstetric complications, and evidence of organ dysfunction. This approach is consistent with the maternal sepsis guidance cited in this review [8,9,11]. However, the included evidence does not establish the diagnostic superiority of any particular monitoring schedule or combination of repeated measurements.

The findings also have implications for settings with limited resources. Community algorithms and shock index studies demonstrate that structured physiological assessment can be implemented in such environments [22,27,30]. Implementation alone does not establish diagnostic accuracy or improved outcomes. Reliable measurement, staff training, referral pathways, and access to further assessment remain necessary, and the effectiveness of these approaches requires prospective evaluation.

A strength of this review was the separation of direct postpartum diagnostic evidence from supportive mixed obstetric and prognostic evidence. This distinction reduced the risk of treating infection detection and severity prediction as equivalent questions. Interpretation also considered measurement timing, reference standards, missing observations, overlapping cohorts, and methodological quality. The decision to avoid pooling preserved clinically important differences, although it limited the ability to provide a summary estimate.

Limitations of the review process include the absence of prospective registration, restriction to English-language reports, failure to retrieve 11 potentially relevant reports, and lack of author contact for missing information. Supplementary searching was less reproducible than the structured database searches. Publication bias and selective reporting bias were not formally assessed, and no formal GRADE assessment was undertaken. These limitations affect the completeness and transparency with which the available

evidence can be evaluated.

The evidence base had additional limitations. Most studies had high risk of bias, only six provided direct postpartum evidence, and several analyses involved small event numbers, selected populations, heterogeneous outcomes, incomplete observations, or limited external validation [17-35]. Many mixed cohorts did not provide separate postpartum estimates. Consequently, confidence in the magnitude and generalizability of reported performance remains limited.

Future studies should recruit consecutive postpartum women and report findings separately for clinically meaningful periods, including the first 24 hours, 24 to 72 hours, four to seven days, and eight to 42 days after childbirth. Diagnostic thresholds and analysis plans should be specified in advance, and reference standards should be independent and clinically adjudicated. Reporting should include complete 2 by 2 data, confidence intervals, and transparent handling of missing observations. Prognostic models should additionally report calibration and undergo external validation. Direct comparisons of repeated measurements, changes from baseline, and fixed thresholds would help determine whether these approaches improve recognition and clinical outcomes.

Conclusions

Routine vital signs remain essential for postpartum surveillance, but no isolated sign or existing warning score can currently be relied upon to confirm or exclude postpartum maternal sepsis. Repeated complete observations, attention to physiological trends, and interpretation within the postpartum clinical context are more appropriate than reliance on a single threshold. Prospective postpartum-specific studies using independent reference standards and external validation are required before any existing score can be recommended as a stand-alone diagnostic method.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Asim Ahmed, Reem E. Elshaikh, Zainab Alshami, Haytham Mohammed Osman Altigani, Thouraya Mohamed Ahmed Elawad

Acquisition, analysis, or interpretation of data: Asim Ahmed, Reem E. Elshaikh, Zainab Alshami, Haytham Mohammed Osman Altigani, Thouraya Mohamed Ahmed Elawad

Drafting of the manuscript: Asim Ahmed, Reem E. Elshaikh, Thouraya Mohamed Ahmed Elawad

Critical review of the manuscript for important intellectual content: Asim Ahmed, Reem E. Elshaikh, Zainab Alshami, Haytham Mohammed Osman Altigani, Thouraya Mohamed Ahmed Elawad

Supervision: Asim Ahmed, Reem E. Elshaikh

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Cresswell JA, Alexander M, Chong MY, et al.: Global and regional causes of maternal deaths 2009-20: a WHO systematic analysis. *Lancet Glob Health*. 2025, 13:e626-34. [10.1016/S2214-109X\(24\)00560-6](https://doi.org/10.1016/S2214-109X(24)00560-6)
2. Acosta CD, Kurinczuk JJ, Lucas DN, Tuffnell DJ, Sellers S, Knight M: Severe maternal sepsis in the UK, 2011-2012: a national case-control study. *PLoS Med*. 2014, 11:e1001672. [10.1371/journal.pmed.1001672](https://doi.org/10.1371/journal.pmed.1001672)
3. Kaukonen KM, Bailey M, Pilcher D, Cooper DJ, Bellomo R: Systemic inflammatory response syndrome criteria in defining severe sepsis. *N Engl J Med*. 2015, 372:1629-38. [10.1056/NEJMoa1415236](https://doi.org/10.1056/NEJMoa1415236)
4. Nakamura E, Naruse K, Miyake R, et al.: Early warning systems to predict severe complications during preterm and postpartum: a scoping review. *J Obstet Gynaecol Res*. 2025, 51:e70079. [10.1111/jog.70079](https://doi.org/10.1111/jog.70079)
5. Singer M, Deutschman CS, Seymour CW, et al.: The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA*. 2016, 315:801-10. [10.1001/jama.2016.0287](https://doi.org/10.1001/jama.2016.0287)
6. Shields LE, Wiesner S, Klein C, Pelletreau B, Hedriana HL: Use of Maternal Early Warning Trigger Tool reduces maternal morbidity. *Am J Obstet Gynecol*. 2016, 214:527.e1-6. [10.1016/j.ajog.2016.01.154](https://doi.org/10.1016/j.ajog.2016.01.154)
7. Bonet M, Nogueira Pileggi V, Rijken MJ, Coomarasamy A, Lissauer D, Souza JP, Gülmezoglu AM: Towards a

- consensus definition of maternal sepsis: results of a systematic review and expert consultation. *Reprod Health*. 2017, 14:67. [10.1186/s12978-017-0321-6](https://doi.org/10.1186/s12978-017-0321-6)
8. Burlinson CE, Sirounis D, Walley KR, Chau A: Sepsis in pregnancy and the puerperium. *Int J Obstet Anesth*. 2018, 36:96-107. [10.1016/j.ijoa.2018.04.010](https://doi.org/10.1016/j.ijoa.2018.04.010)
9. Plante LA, Pacheco LD, Louis JM: SMFM Consult Series #47: sepsis during pregnancy and the puerperium. *Am J Obstet Gynecol*. 2019, 220:B2-10. [10.1016/j.ajog.2019.01.216](https://doi.org/10.1016/j.ajog.2019.01.216)
10. Frequency and management of maternal infection in health facilities in 52 countries (GLOSS): a 1-week inception cohort study. *Lancet Glob Health*. 2020, 8:e661-71. [10.1016/S2214-109X\(20\)30109-1](https://doi.org/10.1016/S2214-109X(20)30109-1)
11. Shields AD, Plante LA, Pacheco LD, Louis JM: Society for Maternal-Fetal Medicine Consult Series #67: maternal sepsis. *Am J Obstet Gynecol*. 2023, 229:B2-19. [10.1016/j.ajog.2023.05.019](https://doi.org/10.1016/j.ajog.2023.05.019)
12. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021, 372:n71. [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)
13. Campbell M, McKenzie JE, Sowden A, et al.: Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ*. 2020, 368:l6890. [10.1136/bmj.l6890](https://doi.org/10.1136/bmj.l6890)
14. Whiting PF, Rutjes AW, Westwood ME, et al.: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011, 155:529-36. [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009)
15. Wolff RF, Moons KG, Riley RD, et al.: PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019, 170:51-8. [10.7326/M18-1376](https://doi.org/10.7326/M18-1376)
16. JBI Manual for Evidence Synthesis. Aromataris E, Munn Z (ed): JBI, Adelaide; 2020. [10.46658/JBIMES-20-01](https://doi.org/10.46658/JBIMES-20-01)
17. Shafik S, Mallick S, Fogel J, Tetrokalashvili M, Hsu CD: The utility of systemic inflammatory response syndrome (SIRS) for diagnosing sepsis in the immediate postpartum period. *J Infect Public Health*. 2019, 12:799-802. [10.1016/j.jiph.2019.04.003](https://doi.org/10.1016/j.jiph.2019.04.003)
18. Hannola K, Hoppu S, Mennander S, Huhtala H, Laivuori H, Tihtonen K: Obstetric early warning system to predict maternal morbidity of pre-eclampsia, postpartum hemorrhage and infection after birth in high-risk women: a prospective cohort study. *Midwifery*. 2021, 99:103015. [10.1016/j.midw.2021.103015](https://doi.org/10.1016/j.midw.2021.103015)
19. Bakhtawar S, Sheikh S, Qureshi R, et al.: Risk factors for postpartum sepsis: a nested case-control study. *BMC Pregnancy Childbirth*. 2020, 20:297. [10.1186/s12884-020-02991-z](https://doi.org/10.1186/s12884-020-02991-z)
20. Ngonzi J, Bebell LM, Fajardo Y, et al.: Incidence of postpartum infection, outcomes and associated risk factors at Mbarara regional referral hospital in Uganda. *BMC Pregnancy Childbirth*. 2018, 18:270. [10.1186/s12884-018-1891-1](https://doi.org/10.1186/s12884-018-1891-1)
21. Ngunyi YL, Halle-Ekane G, Tendongfor N, et al.: Determinants and aetiologies of postpartum pyrexia; a retrospective analysis in a tertiary health facility in the Littoral Region of Cameroon. *BMC Pregnancy Childbirth*. 2020, 20:167. [10.1186/s12884-020-02867-2](https://doi.org/10.1186/s12884-020-02867-2)
22. LeFevre AE, Mir F, Mitra DK, et al.: Validation of community health worker identification of maternal puerperal sepsis using a clinical diagnostic algorithm in Bangladesh and Pakistan. *J Glob Health*. 2021, 11:04039. [10.7189/jogh.11.04039](https://doi.org/10.7189/jogh.11.04039)
23. Champagne HA, Garabedian MJ: Routine screening for sepsis in an obstetric population: evaluation of an improvement project. *Perm J*. 2020, 24:1-10. [10.7812/TPP/19.232](https://doi.org/10.7812/TPP/19.232)
24. Bauer ME, Housey M, Bauer ST, et al.: Risk factors, etiologies, and screening tools for sepsis in pregnant women: a multicenter case-control study. *Anesth Analg*. 2019, 129:1613-20. [10.1213/ANE.0000000000003709](https://doi.org/10.1213/ANE.0000000000003709)
25. Arnolds DE, Carey KA, Braginsky L, Holt R, Edelson DP, Scavone BM, Churpek M: Comparison of early warning scores for predicting clinical deterioration and infection in obstetric patients. *BMC Pregnancy Childbirth*. 2022, 22:295. [10.1186/s12884-022-04631-0](https://doi.org/10.1186/s12884-022-04631-0)
26. Chimwaza Y, Hunt A, Oliveira-Ciabati L, et al.: Early warning systems for identifying severe maternal outcomes: findings from the WHO global maternal sepsis study. *EclinicalMedicine*. 2025, 79:102981. [10.1016/j.eclinm.2024.102981](https://doi.org/10.1016/j.eclinm.2024.102981)
27. Nathan HL, Seed PT, Hezelgrave NL, et al.: Shock index thresholds to predict adverse outcomes in maternal hemorrhage and sepsis: a prospective cohort study. *Acta Obstet Gynecol Scand*. 2019, 98:1178-86. [10.1111/aogs.13626](https://doi.org/10.1111/aogs.13626)
28. Agarwal R, Yadav RK, Mohta M, Sikka M, Radhakrishnan G: Sepsis in Obstetrics Score (SOS) utility and validation for triaging patients with obstetric sepsis in the emergency department: evidence from a low income health care setting. *Obstet Med*. 2019, 12:90-6. [10.1177/1753495X18761522](https://doi.org/10.1177/1753495X18761522)
29. Reicher L, Attali E, Dominski O, et al.: Estimating predictors of severity of group A Streptococcus infection in pregnancy. *J Matern Fetal Neonatal Med*. 2023, 36:2196363. [10.1080/14767058.2023.2196363](https://doi.org/10.1080/14767058.2023.2196363)
30. Kuhrt K, Janneh F, Cole R, et al.: Evaluating shock index for prediction of adverse maternal outcomes related to postpartum haemorrhage and maternal sepsis in Sierra Leone: a prospective observational cohort study. *EclinicalMedicine*. 2025, 89:103578. [10.1016/j.eclinm.2025.103578](https://doi.org/10.1016/j.eclinm.2025.103578)
31. Dohare R, Roy S, Arya A, Parmar S: Sepsis in obstetrics score: predicting morbidity among antenatal and postnatal women. *Bioinformation*. 2025, 21:2090-5. [10.6026/973206300212090](https://doi.org/10.6026/973206300212090)
32. Abutheraa N, Grant J, Mullen AB: Sepsis scoring systems and use of the Sepsis six care bundle in maternity hospitals. *BMC Pregnancy Childbirth*. 2021, 21:524. [10.1186/s12884-021-03921-3](https://doi.org/10.1186/s12884-021-03921-3)
33. Karatas E, Ozkavak OO, Orhan N, et al.: Prognostic value of the obstetric specific scoring systems and four years' experience of a tertiary center. *BMC Pregnancy Childbirth*. 2025, 25:1319. [10.1186/s12884-025-08457-4](https://doi.org/10.1186/s12884-025-08457-4)
34. Ornaghi S, Maraschini A, Buoncristiano M, et al.: Maternal sepsis in Italy: a prospective, population-based cohort and nested case-control study. *Microorganisms*. 2023, 11:105. [10.3390/microorganisms11010105](https://doi.org/10.3390/microorganisms11010105)
35. Kvalvik SA, Zakariassen SB, Overrein S, Rasmussen S, Skrede S, Baghestan E: Obstetric infections and clinical characteristics of maternal sepsis: a hospital-based retrospective cohort study. *Sci Rep*. 2024, 14:6067. [10.1038/s41598-024-56486-4](https://doi.org/10.1038/s41598-024-56486-4)