

SYSTEMATIC REVIEW



Prognostic accuracy of SOFA and qSOFA for mortality among children with infection: a meta-analysis

Zhili Wang^{1,3}, Yu He^{1,3}, Xiaolong Zhang² and Zhengxiu Luo¹✉

© The Author(s), under exclusive licence to the International Pediatric Research Foundation, Inc 2022

BACKGROUND: Age-adjusted Sequential Organ Failure Assessment (SOFA) and age-adjusted quick SOFA (qSOFA) scores have been developed to predict poor outcomes in children with infection. We investigated the prognostic performance of age-adjusted SOFA and age-adjusted qSOFA scores and compared them with the systemic inflammatory response syndrome (SIRS) criteria for predicting mortality in children with infection.

METHODS: A bivariate random-effects regression model was used for synthesis of diagnostic test data.

RESULTS: A total of 14 studies involving 70,194 participants were included. The pooled sensitivity for age-adjusted SOFA, age-adjusted qSOFA, and SIRS were 0.82 (95% CI, 0.74–0.88), 0.46 (95% CI, 0.22–0.71), and 0.79 (95% CI, 0.66–0.88), respectively. The pooled specificity for age-adjusted SOFA, age-adjusted qSOFA, and SIRS were 0.62 (95% CI, 0.45–0.77), 0.90 (95% CI, 0.66–0.98), and 0.39 (95% CI, 0.26–0.54), respectively. The area under the summary receiver operating characteristic curve (AUSROC) for age-adjusted SOFA, age-adjusted qSOFA, and SIRS were 0.82 (95% CI, 0.79–0.85), 0.66 (95% CI, 0.62–0.70), and 0.64 (95% CI, 0.60–0.68), respectively. Different baseline populations, different SOFA adaptation methods and different cut-offs used for age-adjusted SOFA may be potential sources of heterogeneity.

CONCLUSIONS: Age-adjusted SOFA score is a useful tool for predicting mortality in children with sepsis/suspected sepsis.

Pediatric Research (2023) 93:763–771; <https://doi.org/10.1038/s41390-022-02213-6>

IMPACT:

- First study to investigate the prognostic performance of age-adjusted sequential organ failure assessment (SOFA) and age-adjusted quick SOFA (qSOFA) scores in comparison to the systemic inflammatory response criteria (SIRS) for the prediction of mortality in children with sepsis.
- The age-adjusted SOFA score predicts poor outcomes with high sensitivity in children with sepsis
- Low sensitivity limits the utility of age-adjusted qSOFA as a simple predictive tool for adverse outcomes. Developing another enhanced or modified bedside tool with higher sensitivity may be necessary.

INTRODUCTION

Despite the advances in medical care, infectious diseases remain a major disease burden worldwide, especially in children aged five years and younger.¹ Sepsis, defined as a life-threatening syndrome caused by dysregulated host immune response to infection, is a principal cause of morbidity and mortality in children.^{2,3} Globally, an estimated incidence of pediatric sepsis was roughly 48 per 100,000 person-years (1.2 million cases per year).⁴ Mortality for children with sepsis ranges from 4% to 50%, depending on severity of illness, medical resources, and geographic location.⁵

Early recognition and early initiation of treatment are critical to improving outcomes for patients with sepsis.⁶ Systemic inflammatory response syndrome (SIRS) criteria are being used to screen children with infection for development of sepsis since 2005.⁷ However, SIRS have been criticized as screening and prognostic

tools because of inadequate sensitivity and specificity (sensitivity ranged from 0.21 to 0.83, specificity from 0.29 to 0.62).^{8–10} Recently, the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) Task Force replaced the SIRS criteria with the new Sequential (Sepsis-related) Organ Failure Assessment (SOFA) scoring system.¹¹ The SOFA score was used to evaluate the severity of organ dysfunction in adult patients with infection.^{11,12} In addition, the concept of the quick SOFA (qSOFA) was suggested as a simple bedside tool to identify patients with sepsis outside the intensive care units (ICUs), when no laboratory tests are available.^{11,12}

Regrettably, neither SOFA nor qSOFA were developed for children, implying an urgent need for pediatric scores. To fill the gap, researchers have attempted to develop age-adjusted SOFA and age-adjusted qSOFA scores for pediatric patients by using age-specific variables.^{10,13} Recently, several studies have been

¹Department of Respiratory Medicine Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Key Laboratory of Pediatrics, Chongqing 400014, China. ²Department of Pediatrics, Jiangjin District Central Hospital, Chongqing 400014, China. ³These authors contributed equally: Zhili Wang, Yu He. ✉email: luozhengxiu816@hospital.cqmu.edu.cn

Received: 30 October 2021 Revised: 14 June 2022 Accepted: 27 June 2022
Published online: 28 July 2022

conducted to validate the age-adjusted SOFA and age-adjusted qSOFA criteria by assessing their performance in the identification of pediatric patients with confirmed or suspected infection who are likely to have adverse outcomes.^{14–17} However, these studies have produced conflicting evidence regarding predictive accuracy of age-adjusted SOFA and age-adjusted qSOFA scores. It is currently unclear whether age-adjusted SOFA and age-adjusted qSOFA scores have prognostic value for unfavorable outcomes in children with infection.

Thus, the purpose of this study was to evaluate the prognostic value of age-adjusted SOFA and age-adjusted qSOFA and compare them with the SIRS criteria for predicting mortality in children with infection in the hospital setting.

METHODS

This meta-analysis is reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines.¹⁸ The study protocol was registered with the PROSPERO (CRD42021232441).

Data sources and search strategy

We searched Pubmed, Web of science, Cochrane library, and Embase databases from inception to July 10, 2021. The detailed search strategy is presented in Supplementary file 1. We did not restrict our search by year of publication. Only English-language articles were included. For gray literature search, we manually searched the bibliographies of all included studies to identify potentially eligible studies.

Study selection and outcomes

Studies were included based on the following inclusion criteria: (1) involved children (aged 29 days–18 years) with suspected or confirmed infection in the ED, hospital wards, or the ICU; (2) evaluated the age-adjusted SOFA, age-adjusted qSOFA scores or the SIRS criteria for predicting mortality; (3) reported enough data to formulate 2×2 table (true positive (TP), false positive (FP), true negative (TN), and false negative (FN)) for further analysis; (4) performed retrospectively or prospectively and published as full-length articles in peer-reviewed journals.

Articles were excluded: (1) if they were case reports, case series, abstracts or letters; (2) if there were insufficient data to calculate the outcome data (TP, FP, TN, FN).

The primary outcome was in-hospital mortality. For studies in which researchers did not investigate in-hospital mortality, we used the 28- or 30-day mortality instead. The primary outcome measures were the sensitivity and specificity; secondary measures included the positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR), and area under the summary receiver operating characteristic curve (AUSROC).

Data extraction, definitions, and quality assessment

Two reviewers (ZW, YH) independently screened the titles and abstracts and retrieved full-text articles for potentially relevant references. All selections were decided by consensus.

Two investigators (ZW and YH) independently collected the following variables from the included articles: authors, year of publication, country, study design, study setting, cut-off, mean or median age, number of patients included, criteria to diagnosis of infection, variables used for developing age-adjusted SOFA and

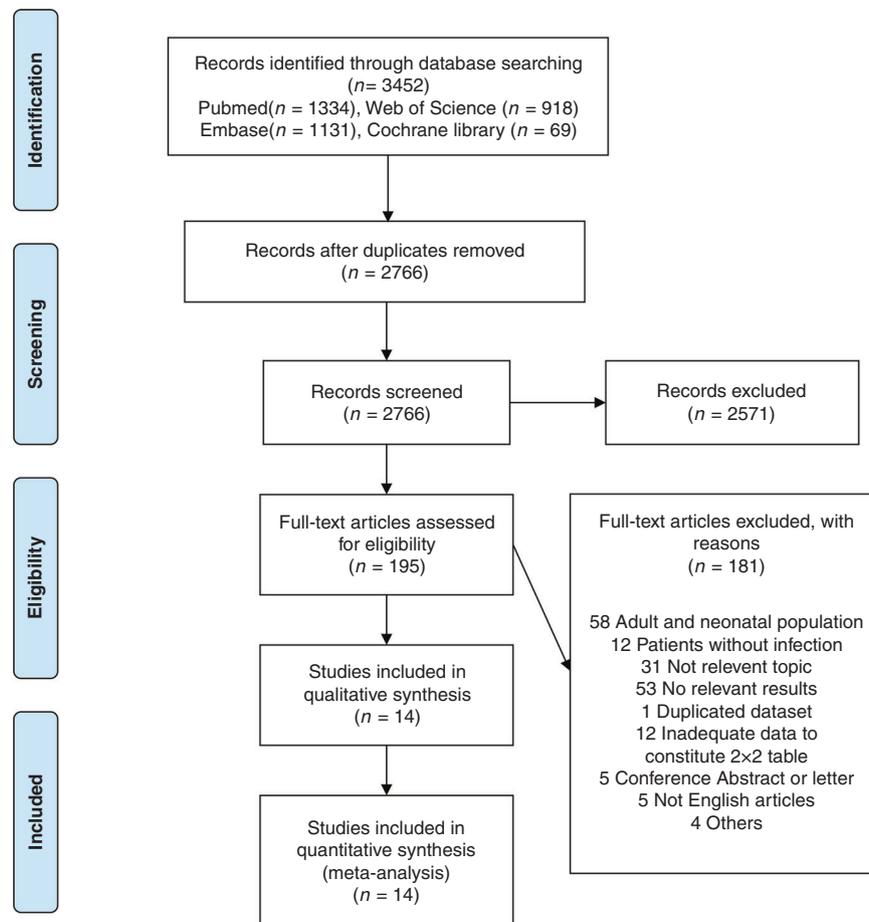


Fig. 1 Flow diagram of study inclusion. PRISMA flow diagram for literature search and screening.

Table 1. Characteristics of studies included in the meta-analysis.

Author, year [reference]	Design	Country	Study size	Age (median)	Mortality rate (%)	Location	Infection Criteria	Index test [reference]	Cut-off values used for developing age-specific qSOFA and SOFA scores	Cut-off	Timing of the score measurement	Measured mortality
Wiens et al. 2016 [31]	Prospective, multicenter	Uganda	1,307	18 months	5.0	Pediatric ward	Proven or suspected infection	SIRS	NA	≥ 2	NR	In-hospital mortality
Kortz et al. 2017 [28]	Prospective, single-center	United States	402	25 months	14.2	ED	Diagnostic or sepsis (SIRS ≥ 2)	SIRS	NA	≥ 3	NR	In-hospital mortality
Schlapbach et al. 2018 [10]	Prospective, multicenter	Australia, New Zealand	2,594	13 years	5.8	ICU	Suspected or proven infection	SIRS, age-adapted SOFA and qSOFA proposed by Schlapbach et al. [10]	Cardiovascular and renal dysfunction cut-offs from PELOD-2 were used for developing age-specific SOFA. Age-specific qSOFA scores were defined by applying age-specific cut-offs for RR and SBP, respectively, as per the 2005 Pediatric Sepsis definitions	≥ 2 for all criteria	Upon ICU admission	In-hospital mortality
Nassau et al. 2018 [35]	Retrospective, single-center	Netherlands	864	NR	0.7	ED	Suspected bacterial infection	SIRS, age-adapted qSOFA proposed by Schlapbach et al. [10]	Age-specific cut-offs for RR and SBP were modified using cut-offs from the 2005 Pediatric Sepsis definitions	≥ 2 for all criteria	Within 24 h after ED entry	30-day mortality
Peters et al. 2018 [17]	Retrospective, single-center	Canada	40,228	NR	4.2	ICU	Discharge diagnosis of infection	Age-adapted qSOFA proposed by Schlapbach et al. [10]	Age-specific cut-offs for RR and SBP were modified using cut-offs from the 2005 Pediatric Sepsis definitions	≥ 2	Within 12 h of ICU admission	In-hospital mortality
Nariadhara et al. 2019 [9]	Prospective, single-center	Tanzania	1,030	15 months	6.3	ED	SIRS ≥ 2	SIRS	NA	≥ 2	Upon enrollment to the study	In-hospital mortality
Wu et al. 2019 [32]	Retrospective, single-center	China	1,831	12.8 months	9.0	ICU	Existent or suspected infection	SIRS, age-adapted SOFA proposed by Schlapbach et al. [10]	Age-specific SOFA was modified by using cardiovascular and renal dysfunction cut-offs from PELOD-2	≥ 2 for all criteria	Within 24 h after ICU admission	In-hospital mortality
Wulandari et al. 2019 [15]	Prospective, single-center	Indonesia	30	2.2 months	46.7	ICU and HCU	Diagnostic of sepsis (SIRS ≥ 2)	SIRS, age-adapted SOFA proposed by Matics et al. [13]	Cardiovascular and renal variables form the original version of SOFA were modified using cut-offs from PELOD-2; the SpO ₂ /FiO ₂ ratio was used as an alternative surrogate of respiratory dysfunction	≥ 2 for all criteria	Within 24 h after ICU admission	In-hospital mortality
Li et al. 2020 [30]	Retrospective, single-center	China	146	2.7 months	17.1	NR	<i>Klebsiella pneumoniae</i> bloodstream infection	Age-adapted SOFA proposed by Matics et al. [13]	Cardiovascular and renal variables form the original version of SOFA were modified using cut-offs from PELOD-2; the SpO ₂ /FiO ₂ ratio was used as an alternative surrogate of respiratory dysfunction	NR	Within 24 h after the blood culture	30-day mortality

Table 1. continued

Author, year [reference]	Design	Country	Study size	Age (median)	Mortality rate (%)	Location	Infection Criteria	Index test [reference]	Cut-off values used for developing age-specific SOFA and qSOFA scores	Cut-off	Timing of the score measurement	Measured mortality
Romaine et al. 2020 [16]	Retrospective, single-center	United Kingdom	12,241	2.5 years	0.04	ED	Febrile children if clinician decided to perform blood tests	Age-adapted qSOFA proposed by Schlapbach et al. [10]	Age-specific cut-offs for RR and SBP were modified using cutoffs from the 2005 Pediatric Sepsis definitions	≥2	Worst value recorded in the ED	Sepsis-related in-hospital mortality
Lalitha et al. 2020 [29]	Prospective, single-center	India	240	36 months	17.5	ICU	Suspected or proven infection with SIRS ≥ 2	Age-adapted SOFA proposed by Matics et al. [13]	Cardiovascular and renal variables were modified using cutoffs from PELOD-2; the SpO ₂ :FiO ₂ ratio was used as an alternative surrogate of respiratory dysfunction	≥8	At the time of admission	In-hospital mortality
Zhong et al. 2020 [33]	Retrospective, single-center	China	516	15.8 months	5.4	ICU	Diagnostic of sepsis (SIRS ≥ 2)	Age-adapted SOFA proposed by Matics et al. [13]	Cardiovascular and renal dysfunction cutoffs from PELOD-2; the SpO ₂ :FiO ₂ ratio was used as an alternative surrogate of respiratory dysfunction	≥5	Worst value in the 24 h	In-hospital mortality
Jhang et al. 2021 [14]	Retrospective, single-center	Korea	135	134 months	18.5	ICU	Diagnostic of septic shock	Age-adapted SOFA proposed by Matics et al. [13]	Cardiovascular and renal variables were modified using cutoffs from PELOD-2; the SpO ₂ :FiO ₂ ratio was used as an alternative surrogate of respiratory dysfunction	NR	Within 24 h after ICU admission	28-day mortality
Liu et al. 2021 [34]	Retrospective, single-center	United States	8,630	6.0 years	2.2	PICU	Suspected infection	SIRS, age-adapted SOFA proposed by Matics et al. [13]	Cardiovascular and renal dysfunction cutoffs from PELOD-2; the SpO ₂ :FiO ₂ ratio was used as an alternative surrogate of respiratory dysfunction	NR	NR	In-hospital mortality

ED Emergency department, FiO₂ Fraction of inspired oxygen, ICU Intensive care unit, MA Not available, NR Not reported, PELOD-2 Pediatric logistic organ dysfunction score-2, qSOFA Quick Sequential (Sepsis-related) Organ Failure Assessment, RR Respiratory rate, Sepsis-3 Third International Consensus Definitions for Sepsis and Septic Shock, SIRS Systemic Inflammatory Response Syndrome, SBP Systolic blood pressure, SOFA Sequential (Sepsis-related) Organ Failure Assessment, SpO₂ Ratio of peripheral oxygen saturation

age-adjusted qSOFA criteria by the individual authors in the included studies, mortality (in-hospital, 28-day, or 30-day), and outcome data (TP, FP, FN, and TN). Discrepancies were resolved by consensus.

The majority of the included studies of the age-adjusted SOFA used the scoring system proposed by Matics et al.¹³, which was termed pediatric SOFA (pSOFA). The pSOFA score was developed by adapting the original SOFA. Specifically, the mean arterial pressure and serum creatinine level cut-offs for the first score of the pediatric logistic organ dysfunction score-2 (PELOD-2)^{19–21} were used to assign a score of 1 in the pSOFA cardiovascular and renal subscores, respectively. The cardiovascular subscores 2 to 4 were kept identical to the original SOFA score. The renal subscores 2 to 4 were modified by increasing the cut-off values for each score by the same factor as the original SOFA score. For the respiratory subscore, the arterial partial pressure of oxygen (PaO₂):fraction of inspired oxygen (FiO₂) ratio cut-offs were kept identical to the original score and the peripheral oxygen saturation (SpO₂):FiO₂ ratio was used as an alternative surrogate of respiratory dysfunction. Additionally, two studies used the another age-adapted SOFA scoring system developed by Schlapbach et al.¹⁰, which is slightly different from the pSOFA score. In brief, the renal and cardiovascular variables of the original SOFA score were modified only using age-specific cut-offs from the PELOD-2. Besides, the respiratory subscore was kept identical to the original SOFA score. On the other hand, to adapt a pediatric version of the qSOFA score, the original tachypnoea and hypotension criteria, based on respiratory rate and systolic blood pressure, were modified using age-specific cut-offs, respectively, as per the 2005 Pediatric Sepsis consensus.^{3,22}

Two investigators (ZW and YH) independently assessed risk of bias of the included studies using the QUADAS-2²³ (Quality Assessment of Diagnostic Accuracy Studies 2) tool by Review Manager (RevMan, version 5.3).

Data synthesis and statistical analysis

A 2 × 2 diagnostic table with the four values, i.e., the TP, FP, FN, and TN was constructed for all studies included. Studies, which did not report these four indexes, relevant values were calculated from the reported specificity, sensitivity, positive predictive value (PPV), and negative predictive value (NPV) in Revman 5.3. Because high between-study heterogeneity was anticipated, the bivariate random-effects regression model²⁴ was used to calculate the pooled sensitivity, specificity, PLR, NLR, and DOR with corresponding 95% credible interval (CI) in STATA 16.0. We also plotted summary receiver operating characteristic curve (SROC) and calculated the AUSROC. Between-study heterogeneity was assessed using the *I*² statistic.²⁵ An *I*² > 50% indicated a substantial heterogeneity. Meta-regression analyses were conducted to explore the potential sources of heterogeneity among included studies with following covariates if data were available: study design (prospective vs. retrospective), disease severity (suspected or confirmed infection vs. sepsis or septic shock), study location (ICU vs. ED or general ward), overall mortality (≥10% vs. <10%), cut-off (predetermined vs. optimal), region, and outcome definitions (in-hospital mortality vs. 28- or 30-day mortality). Deek's funnel plot²⁶ was used to detect publication bias, with *P* < 0.05 indicating potential publication bias.

The overall rating of confidence in pooled prognostic accuracy estimates was assessed using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach.²⁷ Assessments were based on the following criteria: precision, risk of bias of the included studies, consistency, directness of the evidence, and risk of publication bias. Confidence in prognostic accuracy estimates was rated as high, moderate, low, or very low. A GRADE evidence profile was created using the guideline development tool by GRADEpro (gdt.grade.pro.org).

RESULTS

Study search

The PRISMA flow diagram in Fig. 1 shows the literature search process. The initial search identified 3,452 citations. After removing duplicate articles, we screened 2,766 potentially eligible articles. Of these articles, 2,571 were excluded based on title and abstract. A total of 195 articles underwent full-text review. 181 studies were excluded for the reasons presented in Fig. 1. Finally, a total of 14 articles^{9,10,14–17,28–35} (70,194 participants) met our inclusion criteria.

Study characteristics and quality assessment

Demographic variables of the included studies are summarized in Table 1. The number of patients in each study ranged from 30 to 40,228, and the overall mortality rate in each study ranged from 0.04% to 46.7%. Of the included studies, 42.9% were conducted in Asia, 21.4% in North America, 14.3% in Europe, 14.3% in Africa, and 7.1% in Oceania. All studies were observational. Six studies were prospective, and the remaining eight were retrospective. Nine studies evaluated patients at PICU admission,^{10,14,15,17,29,30,32–34} four studies^{9,16,28,35} evaluated patients in the ED, and one study³¹ evaluated patients in general ward. The utility of predicting mortality was evaluated in 4 studies^{10,16,17,35} using age-adjusted qSOFA score and in 8 studies^{9,10,15,28,31,32,34,35} using SIRS criteria. Eight studies^{10,14,15,29,30,32–34} assessed age-adjusted SOFA for mortality prediction, of which six^{14,15,29,30,33,34} employed age-adapted SOFA scoring system proposed by Matics et al.¹³, and two^{10,32} employed scoring system developed by Schlapbach et al.¹⁰ For age-adjusted SOFA, six studies used a pre-specified cut-off value (≥2), one study used a cut-off value of ≥5, and one used a cut-off value of ≥8. The included studies assumed different criteria to define infection in selected patients. Eight studies^{10,17,28,30–32,34,35} considered clinically presumed or suspected infection, four^{14,15,29,33} considered the diagnosis of sepsis or septic shock at ICU admission, one¹⁶ considered the suspicion of infection resulting in obtaining blood tests, and one⁹ considered SIRS-positive patients (SIRS ≥ 2). In terms of outcome measures, 10 studies analyzed in-hospital mortality, 2 evaluated 28-day mortality, and 2 examined 30-day mortality.

Quality assessments using QUADAS-2 criteria are summarized in Fig. 2. The majority of studies have a high or unclear risk of bias in patient selection and index test because of the retrospective nature of their analyses. All included studies were judged as low risk of bias for reference standard, and 9 studies were judged as having low bias in terms of flow and timing. Regarding applicability concerns, all included studies were judged as having low bias in relation to reference standard, whereas the index test do raise questions about applicability in several studies.

Diagnostic accuracy for mortality using age-adjusted SOFA, age-adjusted qSOFA scores and SIRS criteria

Figure 3 shows forest plots and heterogeneity tests of the sensitivity and specificity of age-adjusted SOFA, age-adjusted qSOFA and the SIRS criteria reported in the 14 included studies. Figure 4 shows the SROC curves for three scoring systems in predicting mortality. Summary estimates of all diagnostic accuracy measures are shown in Supplementary file 1: Table S1.

The pooled sensitivity and specificity for age-adjusted SOFA score were 0.82 (95% CI, 0.74–0.88, *I*² = 91.5%) and 0.62 (95% CI, 0.45–0.77, *I*² = 99.6%), respectively. The PLR, NLR, and pooled DOR were 2.18 (95% CI, 1.46–3.25), 0.29 (95% CI, 0.22–0.38), and 7.50 (95% CI, 4.45–12.63), respectively. The pooled sensitivity of the age-adjusted qSOFA score across all included studies was 0.46 (95% CI, 0.22–0.71, *I*² = 97.4%), and the specificity was 0.90 (95% CI, 0.66–0.98, *I*² = 99.9%). The PLR, NLR, and the pooled DOR were 4.63 (95% CI, 1.34–16.02), 0.60 (95% CI, 0.39–0.93), and 7.68 (95% CI, 1.88–31.42), respectively. The pooled sensitivity and specificity for positive SIRS criteria were 0.79 (95% CI, 0.66–0.88, *I*² = 92.4%)

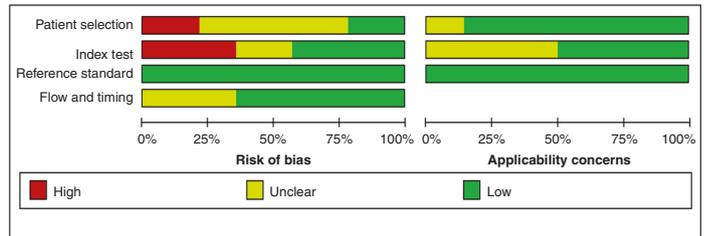
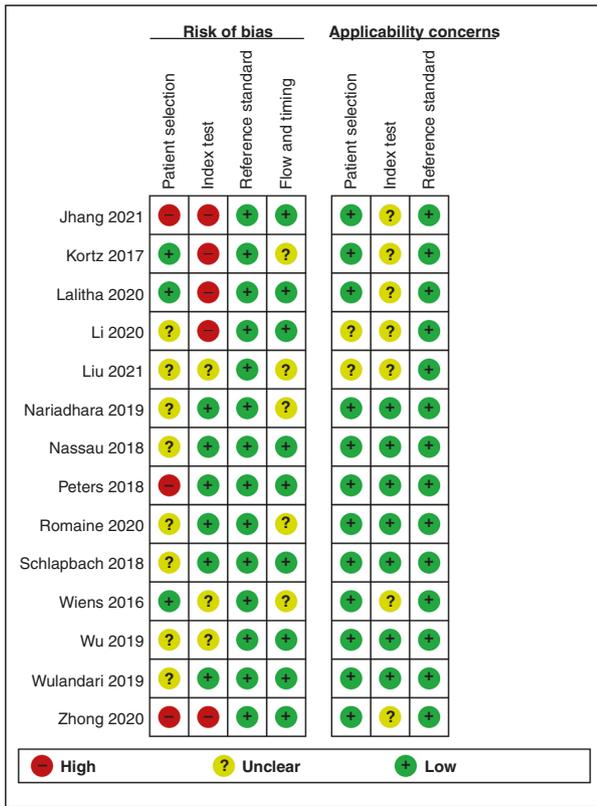


Fig. 2 Methodological quality assessment of included studies based on the QUADAS-2 scale. Assessment of risk of bias for individual studies (left) and risk of bias for studies of overall prognosis (right).

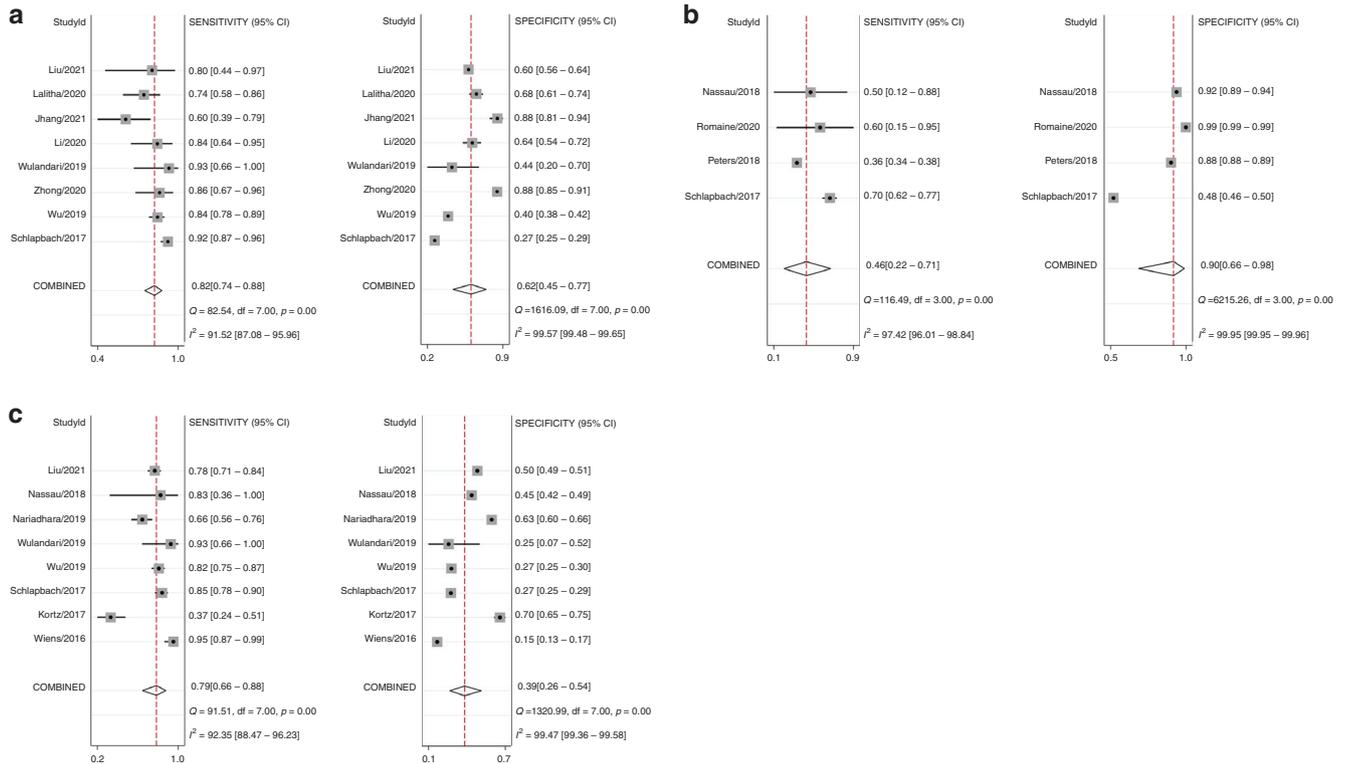


Fig. 3 Forest plots of sensitivity and specificity. a Age-adjusted SOFA, b age-adjusted qSOFA, c SIRS. Point estimates for sensitivity and specificity and 95% confidence intervals are shown with pooled estimates. Q = Cochran Q statistic.

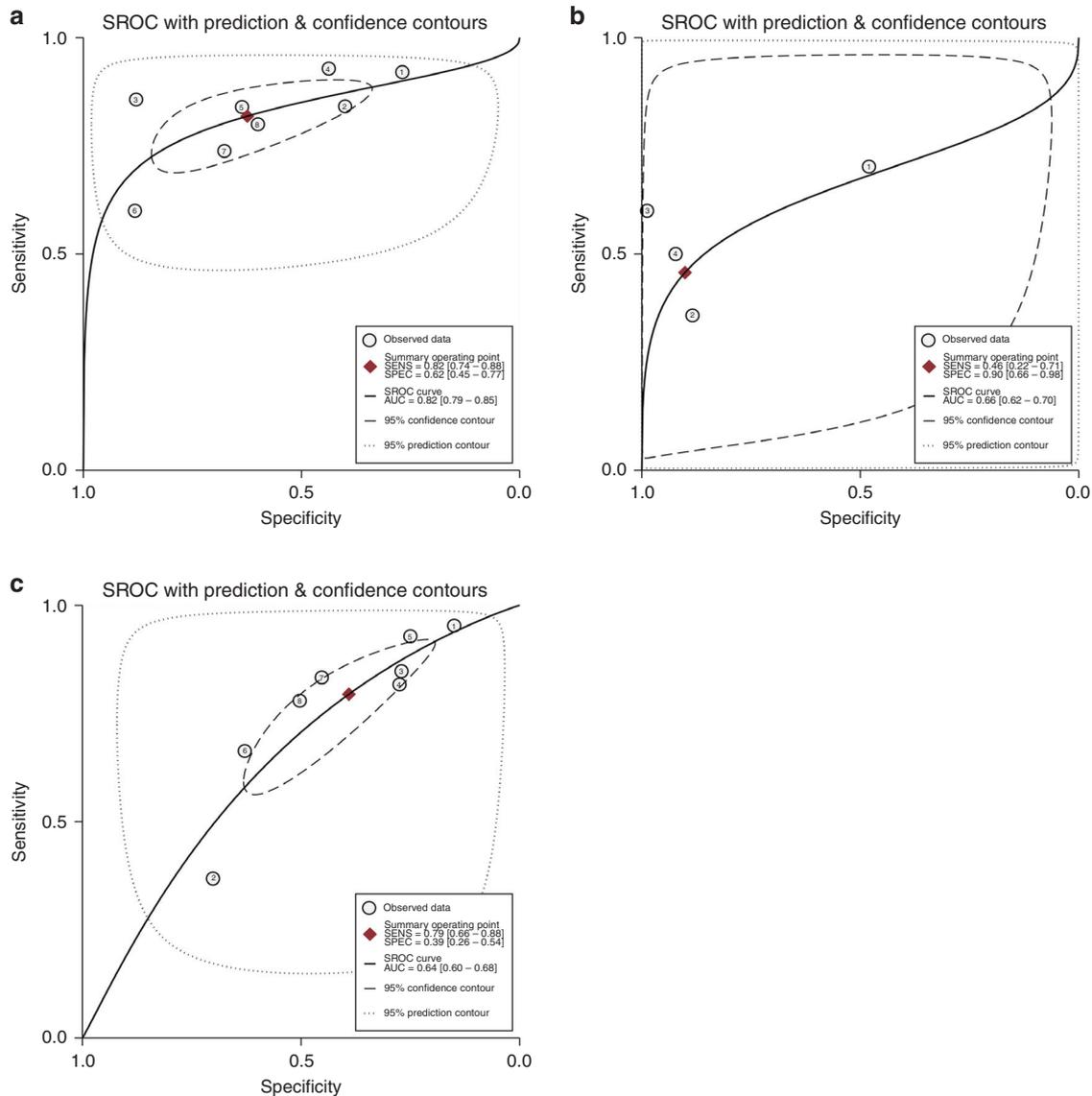


Fig. 4 Summary receiver operating characteristic (SROC) curves for predicting mortality in children with suspected infection. **a** Age-adjusted SOFA, **b** age-adjusted qSOFA, **c** SIRS AUC = area under the curve.

and 0.39 (95% CI, 0.26–0.54, $I^2 = 99.5\%$), respectively. The PLR, NLR, and the pooled DOR were 1.30 (95% CI, 1.15–1.48), 0.53 (95% CI, 0.40–0.69), and 2.48 (95% CI, 1.82–3.38), respectively. The AUSROC and corresponding 95% CI of age-adjusted SOFA, age-adjusted qSOFA, and the SIRS were 0.82 (95% CI, 0.79–0.85), 0.66 (95% CI, 0.62–0.70), and 0.64 (95% CI, 0.60–0.68), respectively.

The GRADE evidence profiles for pooled prognostic accuracy results of age-adjusted SOFA, age-adjusted qSOFA and the SIRS criteria are displayed in Supplementary file 1: Tables S2–4. The overall certainty of evidence for all scoring systems assessed was judged as low to very low due to risk of bias, imprecision and inconsistency.

Meta-regression and subgroup analyses

A high degree of heterogeneity in sensitivity and specificity was present in the included studies (Fig. 3). Univariate meta-regression analysis of sensitivity, and specificity was performed to find potential sources of heterogeneity for age-adjusted SOFA and SIRS (Supplementary file 1: Tables S5–S6) scores. We could not perform further analysis for age-adjusted qSOFA score due to the limited number of studies.

For studies in which researchers evaluated the prognostic accuracy of age-adjusted SOFA score, cut-off value was the probable source of heterogeneity. The pooled specificity for age-adjusted SOFA was higher in studies using an optimal cut-off (0.74 (95% CI 0.57–0.91), $P < 0.01$) than in studies using a predetermined cut-off (0.21 (95% CI 0.06–0.35), $P < 0.01$).

For studies examined prognostic performance of the SIRS criteria, region was the possible source of heterogeneity. The sensitivity of SIRS in North America (0.61 (95% CI 0.36–0.86), $P = 0.03$) was significantly lower than in Africa (0.86 (95% CI 0.70–1.00), $P = 0.73$), Asia (0.85 (95% CI 0.68–1.00), $P = 0.83$), Europe (0.84 (95% CI 0.46–1.00), $P = 0.49$), and Oceania (0.85 (95% CI 0.61–1.00), $P = 0.67$).

Publication bias

Supplementary file 1: Figure S1 shows the assessment of publication bias. Based on the P values of age-adjusted SOFA, age-adjusted qSOFA and the SIRS criteria (0.07, 0.73, and 0.78, respectively) and the corresponding Deek's funnel plot, there was no evidence of publication bias.

DISCUSSION

In the present meta-analysis, we evaluated the prognostic values of age-adjusted SOFA, age-adjusted qSOFA and the SIRS for predicting mortality in children with infection. The certainty of evidence ranged from low to very low for all scoring systems. We found that age-adjusted SOFA is more sensitive and specific than SIRS for mortality prediction. The age-adjusted qSOFA score is more specific, but less sensitive than the SIRS criteria. However, due to the significant heterogeneity among the included studies, our findings should be interpreted with caution.

Our pooled estimates showed that age-adjusted SOFA score had high sensitivity with relatively low specificity. Sepsis is a common cause of multi-organ dysfunction; the SOFA score evaluates the dysfunction of six organ systems, making it more sensitive in predicting poor outcomes.³⁶ The relatively low specificity of the age-adjusted SOFA score indicates that some patients at low risk of mortality may be misclassified as severe. However, early recognition of sepsis and prompt disease management are crucial for decreasing sepsis-related mortality.⁶ High sensitivity is particularly important in identifying critically ill patients. Thus, the age-adjusted SOFA score might provide prognostic value as a useful tool to identify children with infection who are likely to have a poor prognosis. However, there is substantial heterogeneity among the included studies. Instead of using a predetermined cut-off, some studies used an optimal cut-off, which could result in overestimation of prognostic accuracy. Moreover, two age-adapted SOFA scoring system were employed by the included studies, which might be another source of heterogeneity. However, owing to the small number of eligible studies, we could not conduct a further subgroup analysis. Notably, all studies that assessed the prognostic value of age-adjusted SOFA in this review were conducted in the ICU. It remains uncertain whether age-adjusted SOFA could have similar prognostic value in patients outside the ICU. The predictive power of the age-adjusted SOFA score for poor prognosis needs to be further confirmed in different settings. Previous meta-analysis³⁷ on the prognostic accuracy of qSOFA in adult patients showed that the diagnostic specificity of qSOFA score for predicting mortality was higher in studies with overall mortality < 10% than studies with overall mortality \geq 10%. We noticed that mortality rates varied widely between included studies (0.04% to 46.7%), which might have modifier effect for the prognostic accuracy of age-adjusted SOFA. However, our subgroup analysis based on mortality rates did not reveal differences in the prognostic value of age-adjusted SOFA score (Supplementary file 1: Tables S5).

Our results showed that age-adjusted qSOFA had high specificity for mortality prediction but very low sensitivity. There could be several possible reasons for these results. First, hypotension is known to be a very late sign of pediatric sepsis.³⁸ Thus, a predictive tool that incorporates late shock signs, such as hypotension could improve on the specificity but decrease sensitivity. Second, the analyses of prognostic performance for age-adjusted qSOFA were conducted by using data predominantly from ICU patients.^{10,17} There may be differences in characteristics between patients outside the ICU and those in the ICU. For example, respiratory rate and systolic blood pressure in ICU patients may be influenced by mechanical ventilation and vasoactive medications, and mental status may be influenced by sedative agents. This heterogeneity may contribute to a lower sensitivity of the age-adjusted qSOFA for mortality prediction. We attempted to perform subgroup analysis to determine whether there was a difference in between the predictive performance of age-adjusted qSOFA score in ED versus ICU. However, we could not perform subgroup analysis on this subject because of the insufficient number of eligible studies. Further studies are warranted to verify the prognostic and predictive value of age-

adjusted qSOFA in different clinical settings, especially outside the ICU.

Given that the “quick” organ dysfunction criteria are developed as a simple tool to aid physicians in making rapid clinical decisions, the very low sensitivity of the age-adjusted qSOFA score in the identification of sepsis raises concerns about potential delays in recognition and management of sepsis. It may be debatable whether the existing qSOFA parameters are suitable and adequate in predicting poor outcomes for pediatric patients with infection. First, respiratory rate could be affected by multiple non-infectious causes, such as inconvenience and pain. Moreover, fever alone might be a significant determinant of the presence of tachypnea.⁸ Second, as described previously, hypotension is a late sign of pediatric sepsis. Thus, the respiratory rate and blood pressure variables included in pediatric version of qSOFA may be not highly specific for children. Researchers found that heart rate might be superior to respiratory rate in predicting poor outcomes.⁸ Further more, the sensitivity of age-adjusted qSOFA could be improved by substituting the hypotension from the initial version of qSOFA into capillary refill time.¹⁶ More work is needed to determine variables which could be useful in predicting poor outcomes and develop a robust and simple screening tool for pediatric patients.

The SIRS criteria have been used to diagnose pediatric sepsis since 2001, and have shown an inability for poor prognosis prediction for children with infection.⁸ The SIRS criteria were found in >90% of children with fever in the emergency department, but only <5% of those needed intensive care.⁸ Further, a multicenter cohort study found that a positive SIRS criteria were present in over 80% of children admitted to ICU with infection, leading to poor specificity to capture children at high risk for poor prognosis.¹⁰ Consistent with previous studies, we found that although the discriminative ability for mortality using positive SIRS criteria was moderately good, the specificity was very low.

The present study has several strengths. It included a comprehensive search strategy across multiple databases, and multiple subgroup and regression analyses to identify potential sources of bias. The large number of patients included in our study is an additional strength. Moreover, there is a low risk of publication bias and high degree of applicability. Meanwhile, there are several limitations. First, the heterogeneity among the included studies was significant, although heterogeneity is common in meta-analysis of diagnostic studies.³⁹ Second, although we performed multiple meta-regression analyses and the findings of most meta-regression analyses were in keeping with those in the overall study cohort, the results of these analyses should be interpreted with some caution due to lack of statistical power. It's obvious that as we have included patients from different settings, heterogeneity will always exist. Due to a lack of access to individual patient data, it was impossible to adjust the pre-existing clinical risk factors of mortality in these patients. Third, we found only four studies using age-adjusted qSOFA score for predicting mortality. Although our results reveal that the discriminative ability for age-adjusted qSOFA score was higher than SIRS criteria, limited evidence did not allow us to draw a firm conclusion.

CONCLUSION

Age adjusted SOFA score is a useful tool for predicting mortality in children with sepsis/suspected sepsis

DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

REFERENCES

- Liu, L. et al. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the sustainable development goals. *Lancet* **388**, 3027–3035 (2016).
- Plunkett, A. & Tong, J. Sepsis in children. *BMJ* **350**, h3017 (2015).
- Goldstein, B., Giroir, B. & Randolph, A. International consensus conference on pediatric, S. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr. Crit. Care Med.* **6**, 2–8 (2005).
- Fleischmann-Struzek, C. et al. The global burden of paediatric and neonatal sepsis: a systematic review. *Lancet Respiratory Med.* **6**, 223–230 (2018).
- Weiss, S. L. et al. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. *Intensive Care Med.* **46**, 10–67 (2020).
- Seymour, C. W. et al. Time to treatment and mortality during mandated emergency care for sepsis. *N. Engl. J. Med.* **376**, 2235–2244 (2017).
- Carcillo, J. A. & Fields, A. I. Clinical practice parameters for hemodynamic support of pediatric and neonatal patients in septic shock. *Crit. Care Med.* **30**, 1365–1378 (2002).
- Scott, H. F., Deakyn, S. J., Woods, J. M. & Bajaj, L. The prevalence and diagnostic utility of systemic inflammatory response syndrome vital signs in a pediatric emergency department. *Acad. Emerg. Med.* **22**, 381–389 (2015).
- Nariadhara, M. R., Sawe, H. R., Runyon, M. S., Mwafongo, V. & Murray, B. L. Modified systemic inflammatory response syndrome and provider gestalt predicting adverse outcomes in children under 5 years presenting to an urban emergency department of a tertiary hospital in Tanzania. *Trop. Med Health.* **47**, 13 (2019).
- Schlabach, L. J., Straney, L., Bellomo, R., MacLaren, G. & Pilcher, D. Prognostic accuracy of age-adapted sofa, sirs, pelod-2, and Qsofa for in-hospital mortality among children with suspected infection admitted to the intensive care unit. *Intensive Care Med.* **44**, 179–188 (2018).
- Singer, M. et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* **315**, 801–810 (2016).
- Seymour, C. W. et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* **315**, 762–774 (2016).
- Matics, T. J. & Sanchez-Pinto, L. N. Adaptation and validation of a pediatric sequential organ failure assessment score and evaluation of the sepsis-3 definitions in critically ill children. *JAMA Pediatr.* **171**, e172352 (2017).
- Jhang, W. K. & Park, S. J. Evaluation of sepsis-induced coagulopathy in critically ill pediatric patients with septic shock. *Thromb. Haemost.* **121**, 457–463 (2021).
- Wulandari, A., Pudjiastuti, P. & Martuti, S. Severe sepsis criteria, pelod-2, and psofa as predictors of mortality in critically ill children with sepsis. *Paediatrica Indonesiana* **59**, 318–324 (2019).
- Romaine, S. T. et al. Accuracy of a modified qsofa score for predicting critical care admission in febrile children. *Pediatrics* **146**, 4 (2020).
- Peters, C., Murthy, S., Brant, R., Kisson, N. & Gorges, M. Mortality risk using a pediatric quick sequential (sepsis-related) organ failure assessment varies with vital sign thresholds. *Pediatr. Crit. Care Med.* **19**, e394–e402 (2018).
- Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Int. J. Surgery* **8**, 5 (2010).
- Leteurtre, S. et al. Pelod-2: an update of the pediatric logistic organ dysfunction score. *Crit. Care Med.* **41**, 1761–1773 (2013).
- Leteurtre, S. et al. Daily estimation of the severity of organ dysfunctions in critically ill children by using the Pelod-2 score. *Crit. Care* **19**, 324 (2015).
- El-Nawawy, A., Mohsen, A. A., Abdel-Malik, M. & Taman, S. O. Performance of the pediatric logistic organ dysfunction (Pelod) and (Pelod-2) scores in a pediatric intensive care unit of a developing country. *Eur. J. Pediatr.* **176**, 849–855 (2017).
- Gebara, B. M. Values for systolic blood pressure. *Pediatr. Crit. Care Med.* **6** **500**, 500–501 (2005).
- Whiting, P. F. et al. Qadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Intern. Med.* **155**, 529–536 (2011).
- Reitsma, J. B. et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J. Clin. Epidemiol.* **58**, 982–990 (2005).
- Thompson, S. G. Why Sources of heterogeneity in meta-analysis should be investigated. *Bmj* **309**, 1351–1355 (1994).
- Deeks, J. J., Macaskill, P. & Irwig, L. The performance of Tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J. Clin. Epidemiol.* **58**, 882–893 (2005).
- Balshem, H. et al. Grade guidelines: 3. Rating the quality of evidence. *J. Clin. Epidemiol.* **64**, 401–406 (2011).
- Kortz, T. B. et al. Clinical presentation and outcomes among children with sepsis presenting to a public tertiary hospital in Tanzania. *Front Pediatr.* **5**, 278 (2017).
- Lalitha, A. V. et al. Sequential organ failure assessment score as a predictor of outcome in sepsis in pediatric intensive care unit. *J. Pediatr. Intensive Care* **10**, 110–117 (2020).
- Li, S. et al. A risk score based on pediatric sequential organ failure assessment predicts 90-day mortality in children with Klebsiella pneumoniae bloodstream infection. *BMC Infect. Dis.* **20**, 916 (2020).
- Wiens, M. O. et al. Application of sepsis definitions to pediatric patients admitted with suspected infections in Uganda. *Pediatr. Crit. Care Med.* **17**, 400–405 (2016).
- Wu, Z. et al. Accuracy comparison between age-adapted sofa and sirs in predicting in-hospital mortality of infected children at China's Picu. *Shock* **52**, 347–352 (2019).
- Zhong, M. et al. Day-1 Pelod-2 and Day-1 “Quick” Pelod-2 scores in children with sepsis in the Picu. *J. Pediatr. (Rio J.)* **96**, 660–665 (2020).
- Liu, R. et al. Prediction of impending septic shock in children with sepsis. *Crit. Care Explor* **3**, e0442 (2021).
- van Nassau, S. C. et al. Translating sepsis-3 criteria in children: prognostic accuracy of age-adjusted quick sofa score in children visiting the emergency department with suspected bacterial infection. *Front Pediatr.* **6**, 266 (2018).
- Liu, V. et al. Hospital deaths in patients with sepsis from 2 independent cohorts. *JAMA* **312**, 90–92 (2014).
- Song, J. U., Sin, C. K., Park, H. K., Shim, S. R. & Lee, J. Performance of the quick sequential (Sepsis-Related) organ failure assessment score as a prognostic tool in infected patients outside the intensive care unit: a systematic review and meta-analysis. *Crit. Care* **22**, 28 (2018).
- Brierley, J. et al. Clinical practice parameters for hemodynamic support of pediatric and neonatal septic shock: 2007 update from the American college of critical care medicine. *Crit. Care Med.* **37**, 666–688 (2009).
- Dinnes, J., Deeks, J., Kirby, J. & Roderick, P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol. Assess.* **9**, 1–113 (2005).

AUTHOR CONTRIBUTIONS

Z.W. is responsible for the conception, data search, data analysis, statistical analyses, data interpretation, writing and revising the manuscript. Y.H. is responsible for inclusion and exclusion of studies, assessment of methodological quality, data extraction, writing and revising the manuscript. X.Z. is responsible for inclusion and exclusion of studies, assessment of methodological quality, data extraction, and statistical analyses. Z.L. is responsible for the supervision and reviewed the manuscript for important intellectual content. All authors approved the manuscript.

COMPETING INTEREST

The authors declare no competing interests.

CONSENT STATEMENT

Not applicable

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41390-022-02213-6>.

Correspondence and requests for materials should be addressed to Zhengxiu Luo.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.