Diagnostic Accuracy of Neonatal Sepsis Scoring Systems in Preterm Infants

Rudy[1], Mega Hasanul Huda[1], Siti Chodidjah[1]
[1]Faculty of Nursing, Universitas Indonesia, Depok, Indonesia
Correspondence: **Mega Hasanul Huda**: Kampus UI Depok, Indonesia; megahasanulhuda@gmail.com

ABSTRACT

Neonatal sepsis remains a major cause of morbidity and mortality among preterm infants, particularly in low- and middle-income countries. Early identification through clinical scoring systems, such as the Singh Score, Rosenberg Score, and Bekhof Score, may improve diagnostic accuracy and support timely management. This study aimed to systematically evaluate the diagnostic performance of these three clinical scoring instruments in detecting neonatal sepsis. A systematic search was conducted across seven electronic databases from inception to August 2025. Studies assessing the diagnostic accuracy of the selected scoring systems were included, critically appraised, and analyzed. Data on sensitivity, specificity, and other diagnostic accuracy measures were extracted and synthesized. Five eligible studies were identified, most of which were conducted in low- and middle-income countries, with sample sizes ranging from 80 to 658 neonates. The pooled sensitivity of the three instruments ranged from 0.56 to 0.76, with the Singh Score demonstrating the highest sensitivity (76%), followed by the Bekhof (63%) and Rosenberg (56%) scores. The pooled specificity ranged from 0.33 to 0.69, indicating low to moderate ability to exclude non-sepsis cases. Overall, these instruments demonstrate moderate sensitivity but limited specificity and are more suitable for ruling out sepsis rather than definitively confirming the diagnosis. Their use should be integrated with comprehensive clinical assessment and additional laboratory investigations. These scoring systems may aid in risk stratification and help reduce unnecessary treatment by identifying neonates at low risk of sepsis.
Keywords: neonatal sepsis; preterm infants; clinical scoring system; diagnostic accuracy

INTRODUCTION

Neonatal sepsis remains one of the leading causes of morbidity and mortality among newborns and continues to be a major contributor to deaths in children under five years of age. Nearly half (47%) of all under-five deaths occur during the neonatal period (<28 days). It is estimated that approximately 550,000 neonatal deaths annually are caused by infections, including sepsis. Although the global neonatal mortality rate has declined by 12.9%, the incidence of neonatal sepsis has paradoxically increased by 12.8% [1]. This trend underscores the persistent burden of neonatal infections, particularly among vulnerable populations such as preterm infants.

In low- and middle-income countries (LMICs), the burden of neonatal sepsis is even more substantial. Suspected neonatal sepsis has been reported at a rate of 166 per 1,000 live births; however, only 46.9 cases per 1,000 are laboratory-confirmed, with an estimated mortality of 0.83 per 1,000 patient-days [2]. The high mortality rate is frequently associated with delayed diagnosis, limited access to reliable blood culture results, and suboptimal referral systems. These systemic constraints contribute to delayed treatment initiation and poorer clinical outcomes.

Beyond acute mortality, neonatal sepsis is also associated with significant long-term consequences, particularly in preterm infants [3]. Survivors are at increased risk of neurodevelopmental impairments, including visual and hearing disorders, cerebral palsy, and cognitive developmental delays [4]. Furthermore, neonatal sepsis has been linked to serious complications such as necrotizing enterocolitis (7.4%) and chronic lung disease (58.1%) [5]. These complications not only increase healthcare utilization but also impose long-term socioeconomic burdens on families and health systems. Therefore, early recognition and prompt management are essential to reduce both mortality and long-term morbidity, especially in resource-limited settings.

A major challenge in diagnosing neonatal sepsis is the limited availability of diagnostic resources. Approximately 40% of hospitals in LMICs lack facilities for blood culture testing [6], despite blood culture being considered the gold standard for sepsis diagnosis [7]. In response to these limitations, several symptom-based clinical scoring instruments have been developed to support early identification. Singh proposed seven clinical signs, Rosenberg identified five clinical indicators, and Bekhof introduced four clinical parameters to facilitate early diagnosis [8-10]. These instruments are intended to provide practical and rapid assessment tools in settings where laboratory confirmation is not readily available.

However, validation studies have demonstrated considerable variability in the diagnostic performance of these instruments. For example, the Singh instrument demonstrated sensitivity and specificity values of 83% and 32%, respectively, in India, but only 56.6% and 52.1% in Bangladesh [8,9]. The Rosenberg instrument reported values of 77.1% and 50% in Bangladesh, compared with 46% and 72% in South Africa [9,11]. Similarly, the Bekhof instrument showed wide variation, ranging from 97% sensitivity and 37% specificity in the Netherlands to 55% sensitivity and 71% specificity in South Africa [10,11]. These discrepancies highlight inconsistencies in diagnostic performance across different populations and healthcare contexts.

The variability in reported sensitivity and specificity suggests that no single clinical scoring instrument has demonstrated consistently reliable accuracy across diverse settings, and direct comparative evidence remains limited. Therefore, this study aims to systematically review and critically evaluate the diagnostic accuracy of symptom-based clinical scoring instruments for neonatal sepsis, in order to determine their sensitivity, specificity, and overall clinical utility, and to provide evidence-based recommendations for their application in clinical practice, particularly in resource-limited settings.

METHODS

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines and was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO) on August 11, 2025 (registration number CRD420251124655). As this study synthesized data from previously published articles and did not involve direct contact with human participants or access to identifiable individual data, ethical approval was not required.

Statistical analyses were performed using Stata version 17.0 (StataCorp, College Station, TX) with the *midas* command for diagnostic test accuracy meta-analysis. Pooled estimates of sensitivity and specificity were calculated using a bivariate generalized linear mixed model. The literature search was conducted in June 2025 across seven electronic databases: Sage Journal, ScienceDirect, Oxford Academic, PubMed, Wiley Online Library, ProQuest, and Springer. The search strategy combined the following keywords: (preterm infants OR premature baby) AND (sepsis risk score OR infection prediction score OR late-onset sepsis OR predictive clinical score) AND (validation OR diagnostic accuracy), restricted to titles and abstracts.

Eligible studies included cross-sectional and cohort designs evaluating the sensitivity and specificity of clinical symptom-based prediction instruments for neonatal sepsis occurring ≥72 hours after birth, including preterm and low birth weight infants. Articles were screened based on topic relevance, participant characteristics, study design, and index tests. Duplicate and retracted articles were excluded. Full-text articles were independently reviewed by two investigators (R and MH), and discrepancies were resolved through discussion and consensus.

The predictive instruments analyzed comprised several previously developed clinical scores. The Singh instrument initially included 16 clinical signs, later simplified to seven signs, with reported sensitivity and specificity of 83% and 32%, respectively (cut-off ≥2) [8]. The Clinical Score, derived from the Singh instrument, demonstrated sensitivity of 90% and specificity of 22.5% (cut-off ≥1) [12]. The Rosenberg score, based on five clinical signs, reported sensitivity/specificity values of 77.1%/50% and 56.5%/52.1% in different validation studies, using the same cut-off (≥1) [9]. Bekhof developed a monogram based on five clinical indicators without laboratory parameters, yielding sensitivity of 97% and specificity of 37% (cut-off ≥1) [10]. Lloyd compared these three instruments and reported the following sensitivity and specificity values: Singh 74%/33%, Rosenberg 46%/72%, and Bekhof 55%/71% [11].

Data extraction from eligible studies was performed independently by two reviewers (R and MH). Extracted information included the first author's name, year of publication, study location, study design, instrument used, sample size, gestational age, birth weight, cut-off value, area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Any discrepancies in extracted data were discussed until agreement was reached.

Methodological quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool. This instrument evaluates four domains: patient selection, index test, reference standard, and flow and timing. Each domain was rated as "yes," "no," or "unclear." The risk of bias was considered low if all signaling questions within a domain were answered "yes," high if any were answered "no," and unclear if insufficient information was available to permit judgment.

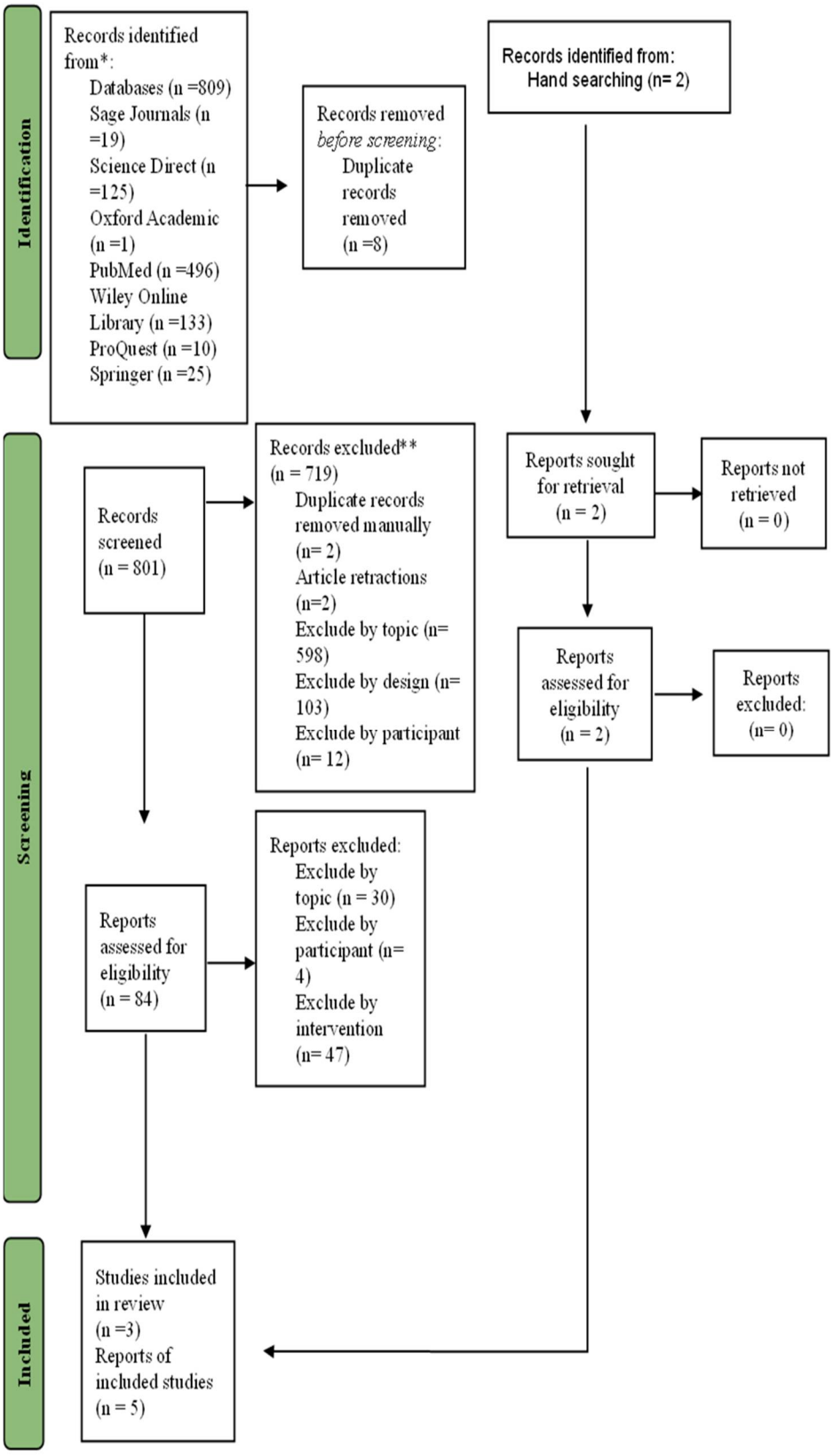

Figure 1. Flow diagram for neonatal sepsis scoring systems

RESULTS

The initial literature search identified a substantial number of records. After the removal of duplicates and screening according to predefined inclusion and exclusion criteria, five primary studies met the eligibility criteria and were included in the final analysis (PRISMA flow diagram) (Figure 1). Among these studies, most evaluated the Singh instrument, whereas the Rosenberg and Bekhof instruments were each reported in two studies. All instruments relied exclusively on simple clinical indicators without requiring complex laboratory investigations, highlighting their potential applicability in resource-limited healthcare settings.

Four articles evaluating the Singh instrument included a total of 1,100 neonates with 1,359 episodes of late-onset sepsis, comprising three cohort studies (India, Bangladesh, and the Netherlands) and one cross-sectional study (South Africa). Two studies assessing the Rosenberg instrument involved 818 neonates with 1,037 episodes, conducted in Bangladesh and South Africa. Similarly, two studies examining the Bekhof instrument included 800 preterm infants with 1,028 episodes, using cohort and cross-sectional designs in the Netherlands and South Africa (Table 1).

Table 1. Characteristics of included studies

Author, years, country	Design	Ms./Ins.	Par./Inc	Sample	GA week (mean, SD)	BW (mean, SD) g	Cof.sc.	AUC	Sen. & Spe. (%)	PPV (%)	NPV (%)
Singh et al., 2003, India [8]	Cohort study	Singh Score	7	Signs: lethargy; tachycardia; hyperthermia; abdominal distension; increased aspirates; retraction; grunting	n: 80 neonates with 105 symptomatic episodes. 91% preterm; 93% low birth weight. Inclusion: babies with any of the 16 signs, onset after 72 h. Exclusion: major congenital malformations	(31.4 ± 3.4)	(1378 ± 507)	≥ 2	NA	83 & 32	33
Rosenberg et al., 2010, Bangladesh [9]	Retrospective	Rosenberg Score	5	Signs: apnoea; hepatomegaly; jaundice; lethargy; pallor	n: 160 with 193 episodes of suspected nosocomial sepsis. Inclusion: age ≤72 h at admission. Exclusion: major congenital malformations; unlikely survival >48 h; early-onset; non-nosocomial; episodes <4 days apart	(30.65)	(1229.2)	≥ 1	0.70	77.1 & 50.0	64.9
		Singh Score	7	Same signs as above				NA	56.6 & 52.1	78.1	28.4
Bekhof et al., 2013, Netherlands [10]	Prospective cohort study	Bekhof Score	4	Signs: increased respiratory support; capillary refill time; pallor/grey skin; central venous catheter	n: 142 with 187 suspected late-onset sepsis episodes. Inclusion: no antibiotics in last 24 h; <34 weeks postconceptional age; >72 h postnatal age; suspicion of infection. Exclusion: major congenital malformations; unlikely survival >48 h; early-onset; non-nosocomial; episodes <4 days apart	(29 w + 6 d ±2+1)	(1207 ±351)	≥ 1	0.84	97 & 37	NA
Kudawla et al., 2007, India [11]	Retrospective	Clinical Score	6	Signs: grunting; abdominal distension; pre-feed aspirates; chest retractions; lethargy; tachycardia; hyperthermia	n: 202 with 220 clinically suspected sepsis episodes. Inclusion: weight 1000–2500 g; clinical suspicion of LONS. Exclusion: major congenital malformations; death <24 h	(31.4 ± 2.5)	(1351 ± 369)	≥ 1	NA	90 & 22.5	30.3
Lloyd et al., 2022, South Africa [12]	Case-control	South African Cohort	Singh Score	N: 658 with 841 infection episodes. Inclusion: 1510 VLBW neonates >72 h. Exclusion: monitoring cultures <72 h; insufficient notes; contaminants	(28.33)	(1062)	≥ 1	0.550	74 & 33	28	78
		Rosenberg Score					≥ 1	0.566	46 & 72	37	79
		Bekhof Score					≥ 1	0.620	55 & 71	40	40

Methodological quality assessment using QUADAS-2 indicated an overall low risk of bias, although some domains were rated as high or unclear (Table 2). In the patient selection domain, one Singh study and Lloyd [11] were rated unclear due to limited recruitment information, and the cross-sectional design contributed to high risk in one study. In the index test domain, Singh and Rosenberg were rated high risk due to inconsistent threshold determination, whereas Bekhof and Kudawla [12] were rated unclear due to incomplete reporting. In the reference standard domain, most studies did not explicitly describe assessment procedures. Nevertheless, regarding applicability, all studies were judged to have low concern and were therefore retained for analysis.

Table 2. Quality assessment of included studies

Study	Patient selection (RoB/App)	Index test (RoB/App)	Reference standard (RoB/App)	Flow and timing (RoB)	Note
Singh [8]	U / L	H / L	U / L	L	RoB = Risk of bias App = Applicability U = Unclear L = Low H = High
Rosenberg [9]	L / L	H / L	U / L	L	
Bekhof [10]	L / L	U / L	U / L	L	
Kudawla [12]	L / L	U / L	L / L	L	
Lloyd [11]	H / L	L / L	U / L	L	

The pooled analysis across instruments (Table 3) demonstrated variability in diagnostic performance with 95% confidence intervals (CI). The Singh instrument showed the highest pooled sensitivity at 0.76 (95% CI: 0.62–0.86), but low specificity at 0.33 (95% CI: 0.24–0.44). In contrast, the Rosenberg instrument demonstrated the lowest sensitivity at 0.56 (95% CI: 0.51–0.61), yet the highest specificity at 0.69 (95% CI: 0.66–0.73). The Bekhof instrument exhibited the most balanced performance, with sensitivity of 0.63 (95% CI: 0.57–0.68) and specificity of 0.65 (95% CI: 0.61–0.68).

Table 3. Pooled sensitivity and specificity

Scale	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)	Note
Singh score [8,9,11,12]	0.76 (0.62 to 0.86)	0.33 (0.24 to 0.44)	CI = Confidence interval
Rosenberg score [9,12]	0.56 (0.51 to 0.61)	0.69 (0.66 to 0.73)	
Bekhof [10,12]	0.63 (0.57 to 0.68)	0.65 (0.61 to 0.68)	

DISCUSSION

The primary findings of this meta-analysis indicate that the Singh instrument demonstrates relatively good sensitivity for detecting suspected cases of neonatal sepsis, although its specificity remains comparatively low. This pattern is consistent with previous studies reporting high sensitivity but limited specificity for the same instrument [8, 12]. When compared with other instruments, the pooled sensitivity of the Singh score was higher than that of the Rosenberg instrument but lower than the findings reported by Kudawla [8,9,11]. In contrast, the Rosenberg instrument tended to show

better specificity, albeit with lower sensitivity [9]. These comparative results suggest that the Singh instrument may be more appropriate as an early screening tool to enhance clinical vigilance, but it is insufficient as a standalone tool to definitively exclude or confirm sepsis.

The variability in sensitivity and specificity across the analyzed studies may be influenced by several factors. Differences in cut-off values, variations in population characteristics (including gestational age, baseline clinical condition, and the number of sepsis episodes), and heterogeneity in the reference standards or diagnostic definitions used across studies likely contributed to the observed discrepancies. Diagnostic definitions varied and included the presence of bacteria detected in blood or cerebrospinal fluid (CSF), positive blood culture results, or clinical suspicion of sepsis accompanied by positive culture findings [8-12]. Such variations in diagnostic criteria can substantially affect estimates of diagnostic accuracy and complicate direct comparisons between studies.

The relatively low specificity observed, particularly for the Singh instrument, indicates that it should not be used as the sole basis for establishing a diagnosis of neonatal sepsis. Instead, its role is more appropriately positioned within an early screening strategy to prompt timely clinical evaluation and management, including the initiation of empirical antibiotic therapy when indicated [13]. However, the literature also emphasizes that irrational or prolonged antibiotic use in neonates may increase the risk of complications, including necrotizing enterocolitis, potentially due to alterations in gut microbiota [14,15]. Therefore, discontinuation of antibiotics when sepsis is not confirmed is strongly recommended as part of global efforts to curb antimicrobial resistance [16]. In the context of diagnostic accuracy testing, high sensitivity supports the function of an instrument as a rule-out or early detection tool rather than a confirmatory diagnostic test. Moreover, estimates of sensitivity and specificity may also be influenced by disease prevalence, which can result in high sensitivity accompanied by low specificity [17]. Notably, most of the primary studies included in this analysis applied a cut-off value of ≥1, which theoretically increases sensitivity at the expense of reduced specificity. This trade-off has been consistently reported in other studies, where improvements in sensitivity are almost invariably associated with declines in specificity estimates [18].

Study design selection may further influence the interpretation of diagnostic performance. Cohort studies are generally considered more robust in assessing temporal relationships and causal inference, although they remain susceptible to bias due to loss to follow-up. Conversely, cross-sectional studies are more feasible and efficient but are limited in their ability to establish temporal relationships between exposure and outcome [19]. Both diagnostic accuracy assessments and comparisons of diagnostic instruments may legitimately employ observational study designs [20]. Interestingly, methodological reviews suggest that findings from randomized controlled trials and observational studies do not necessarily differ significantly in estimating causal associations, provided that the research question is comparable and the risk of bias is adequately controlled [21]. Accordingly, the inclusion of both cohort and cross-sectional designs in this review remains aligned with the study objectives, although methodological heterogeneity should be carefully considered when interpreting the pooled results.

This study also found that the Rosenberg and Bekhof scores tended to demonstrate a more balanced profile between sensitivity and specificity, with the Rosenberg score showing relatively stronger specificity in pooled analysis. Nevertheless, these findings should be interpreted with caution. The limited number of included studies increases the risk of statistical instability and reduces the ability to detect and quantify heterogeneity accurately [22]. Undetected heterogeneity may compromise the consistency and reliability of pooled estimates [23]. Therefore, the present findings should be regarded as complementary evidence that supports existing literature rather than as a definitive basis for clinical decision-making [24-26]. Broader external validation studies involving diverse populations and standardized diagnostic criteria are still required to establish more consistent and generalizable estimates of diagnostic performance for these instruments.

CONCLUSION

These findings indicate that the Singh instrument demonstrates reasonably good performance as an early detection tool for neonates suspected of having sepsis. However, it cannot be used as a definitive basis for establishing a diagnosis and does not replace biomarker-based assessments as the diagnostic gold standard. Meanwhile, the Rosenberg and Bekhof instruments require broader external validation to determine the consistency and generalizability of their results. Overall, symptom-based clinical instruments may serve as useful initial screening tools in resource-limited settings, but their application should be complemented by additional diagnostic investigations to ensure accurate and reliable diagnosis.

Ethical consideration, competing interest and source of funding

-This study adhered to established ethical standards for research conduct. As a systematic review and meta-analysis utilizing previously published data without involving direct human participation or identifiable personal information, formal ethical approval was not required. The review protocol was prospectively registered in PROSPERO (CRD420251124655), and the study was conducted in accordance with PRISMA 2020 guidelines to ensure transparency, methodological rigor, and research integrity.
-There is no conflict of interest related to this study.
-Source of funding is authors.

REFERENCES

1. Li J, Shen L, Qian K. Global, regional, and national incidence and mortality of neonatal sepsis and other neonatal infections, 1990–2019. Front Public Health. 2023 Mar 14;11(1):112-118.
2. Milton R, Gillespie D, Dyer C, Taiyari K, Carvalho MJ, Thomson K, et al. Neonatal sepsis and mortality in low-income and middle-income countries from a facility-based birth cohort: An international multisite prospective observational study. Lancet Glob Health. 2022 May;10(5):e661–672.
3. Cai S, Thompson DK, Anderson PJ, Yang JYM. Short- and long-term neurodevelopmental outcomes of very preterm infants with neonatal sepsis: A systematic review and meta-analysis. Children. 2019 Dec 1;6(12):131-138.
4. Ong WJ, Seng JJB, Yap B, He G, Moochhala NA, Ng CL, et al. Impact of neonatal sepsis on neurocognitive outcomes: A systematic review and meta-analysis. BMC Pediatr. 2024 Aug 7;24(1):505-512.
5. Mukhopadhyay S, Puopolo KM, Hansen NI, Lorch SA, DeMauro SB, Greenberg RG, et al. Neurodevelopmental outcomes following neonatal late-onset sepsis and blood culture-negative conditions. Arch Dis Child Fetal Neonatal Ed. 2021 Sep;106(5):467–473.
6. Murless-Collins S, Kawaza K, Salim N, Molyneux EM, Chiume M, Aluvaala J, et al. Blood culture versus antibiotic use for neonatal inpatients in 61 hospitals implementing with the NEST360 Alliance in Kenya, Malawi, Nigeria, and Tanzania: A cross-sectional study. BMC Pediatr. 2023 Nov 15;23(S2):568-578.
7. Özmeral Odabaşı I. Neonatal sepsis. SiSli Etfal Hastanesi Tip Bulteni / The Medical Bulletin of Sisli Hospital. 2020;1(1):112-118.
8. Singh SA, Dutta S, Narang A. Predictive clinical scores for diagnosis of late onset neonatal septicemia. J Trop Pediatr. 2003 Aug;49(4):235–239.

9. Rosenberg RE, Ahmed ANU, Saha SK, Chowdhury MA, Ahmed S, Law PA, et al. Nosocomial sepsis risk score for preterm infants in low-resource settings. J Trop Pediatr. 2010 Apr 1;56(2):82–89.
10. Bekhof J, Reitsma JB, Kok JH, Van Straaten IHLM. Clinical signs to identify late-onset sepsis in preterm infants. Eur J Pediatr. 2013 Apr 28;172(4):501–508.
11. Kudawla M, Dutta S, Narang A. Validation of a clinical score for the diagnosis of late onset neonatal septicemia in babies weighing 1000-2500 g. J Trop Pediatr. 2007 Nov 4;54(1):66–69.
12. Lloyd LG, Dramowski A, Bekker A, Malou N, Ferreyra C, Van Weissenbruch MM. Performance comparison of infection prediction scores in a South African neonatal unit: A retrospective case-control study. Front Pediatr. 2022 Mar 11;10(1):102-108.
13. Kollef MH, Shorr AF, Bassetti M, Timsit JF, Micek ST, Michelson AP, Garnacho-Montero J. Timing of antibiotic therapy in the ICU. Critical Care. 2021 Oct 15;25(1):360.
14. Cuna A, Morowitz MJ, Sampath V. Early antibiotics and risk for necrotizing enterocolitis in premature infants: a narrative review. Frontiers in Pediatrics. 2023 Feb 14;11:1112812.
15. Kiatchoosakun P, Jirapradittha J, Sirikarn P, Laopaiboon M, Pattanittum P, Chandrakachorn W, et al. Early-onset neonatal sepsis and antibiotic use in Northeast Thailand. Am J Perinatol. 2019 Oct 31;36(12):1295–1303.
16. Fischer AM, Mitchell JL, Stanley KC, Javed MJ. A quality improvement project to reduce antibiotic exposure in premature neonates. Hosp Pediatr. 2023 May 1;13(5):435–448.
17. Murad MH, Lin L, Chu H, Hasan B, Alsibai RA, Abbas AS, et al. The association of sensitivity and specificity with disease prevalence: Analysis of 6909 studies of diagnostic test accuracy. Can Med Assoc J. 2023 Jul 17;195(27):E925–E931.
18. Acharya A, Markar SR, Ni M, Hanna GB. Biomarkers of acute appendicitis: systematic review and cost–benefit trade-off analysis. Surgical Endoscopy. 2017 Mar;31(3):1022-1031.
19. Chirico F. Cross-sectional studies in occupational health research: An overview of strengths and limitations. G Ital Psicol Med Lav. 2023;3(3):86-93.
20. Chassé M, Fergusson DA. Diagnostic accuracy studies. Seminars in Nuclear Medicine. 2019 Mar 1;49(2):87-93.
21. Bosdriesz JR, Stel VS, van Diepen M, Meuleman Y, Dekker FW, Zoccali C, Jager KJ. Evidence-based medicine—when observational studies are better than randomized controlled trials. Nephrology. 2020 Oct;25(10):737-743.
22. Ruppar T. Meta-analysis: How to quantify and explain heterogeneity?. European Journal of Cardiovascular Nursing. 2020 Oct;19(7):646-652.
23. Choi GJ, Kang H. Heterogeneity in meta-analyses: An unavoidable challenge worth exploring. Korean Journal of Anesthesiology. 2025 Aug 1;78(4):301-314.
24. Chow N, Gallo L, Busse JW. Evidence-based medicine and precision medicine: complementary approaches to clinical decision-making. Precision Clinical Medicine. 2018 Sep 1;1(2):60-64.
25. Smith CF, Drew S, Ziebland S, Nicholson BD. Understanding the role of GPs' gut feelings in diagnosing cancer in primary care: a systematic review and meta-analysis of existing evidence. British Journal of General Practice. 2020 Aug 24;8(2):82-88.
26. Lai MC, Anagnostou E, Wiznitzer M, Allison C, Baron-Cohen S. Evidence-based support for autistic people across the lifespan: Maximising potential, minimising barriers, and optimising the person–environment fit. The Lancet Neurology. 2020 May 1;19(5):434-451.