E-ISSN 2250-0944
ISSN 2250-1150
doi: 10.48047/ijprt/16.01.469

Research Article

AI-Driven Early Prediction of Neonatal Sepsis Using Clinical and Laboratory Parameters: A Public Health-Oriented Neonatal Risk Assessment Model

Lubna Manzoor[1], Arooj Haider[2], Javaria Younus[3], Tahir Mahmood Butt[4], Zille Huma Mustehsan[5], Shahab Ali[6]

1. Senior Registrar, Pediatrics, Abwa Hospital, Faisalabad, Pakistan.
2. Health Services Academy, Islamabad, Pakistan.
3. Assistant Professor, Neonatology, University of Child Health Sciences, Children's Hospital, Lahore, Pakistan.
4. Associate Professor, Department of Community Medicine, CMH Kharian Cantt Medical College, Kharian, Pakistan.
5. Associate Professor, Department of Community Medicine, NUST School of Health Sciences, National University of Sciences and Technology, Islamabad, Pakistan.
6. Senior Lecturer Community Medicine and Public Health Department, Bahria University College of Medicine Islamabad

Abstract

Background:
Neonatal sepsis remains a major contributor to neonatal morbidity and mortality worldwide, particularly in low- and middle-income countries, where delayed diagnosis and limited neonatal intensive care resources significantly compromise survival outcomes. Early recognition of neonatal sepsis remains challenging because initial clinical manifestations are often subtle and nonspecific. Artificial intelligence (AI)-driven predictive models have recently emerged as promising tools for improving early detection and risk stratification in neonatal care. **Aim:**
To develop and evaluate an AI-driven neonatal risk assessment model using clinical and laboratory parameters for the early prediction of neonatal sepsis in a tertiary care neonatal intensive care unit. **Methodology:**
A prospective diagnostic prediction model study was conducted on 240 neonates admitted to the NICU between January 2023 and December 2024. Demographic characteristics, maternal risk factors, clinical features, and laboratory parameters were prospectively collected and used to develop and internally validate machine learning models, including Logistic Regression, Random Forest, and Support Vector Machine, for early prediction of neonatal sepsis. Model performance was assessed using accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve.
Results: The final AI-based prediction model demonstrated an accuracy of 91.3%, sensitivity of 89.5%, specificity of 92.8%, and an area under the receiver operating characteristic curve of 0.94 for identifying neonatal sepsis. Respiratory distress, temperature instability, elevated C-reactive protein, abnormal leukocyte count, thrombocytopenia, and prolonged rupture of membranes were identified as significant predictors of neonatal sepsis ($p < 0.001$).
Conclusion: Neonates identified as high-risk by the AI model demonstrated significantly higher rates of intensive care requirement and prolonged hospitalization. These findings suggest that AI-assisted neonatal sepsis prediction models may substantially improve early clinical decision-making, optimize neonatal triage, and strengthen neonatal

public health surveillance in resource-constrained healthcare settings.

Keywords: Neonatal sepsis; Artificial intelligence; Risk prediction model; Machine learning; Neonatal intensive care

Introduction

Neonatal sepsis remains one of the leading causes of neonatal mortality and long-term neurodevelopmental impairment worldwide, accounting for a substantial proportion of deaths among infants younger than 28 days of life. The burden of neonatal sepsis is disproportionately higher in low- and middle-income countries because of delayed diagnosis, inadequate infection control practices, limited laboratory facilities, and shortages of neonatal intensive care resources [1-3]. Despite improvements in neonatal care, neonatal sepsis continues to pose major diagnostic and therapeutic challenges because clinical manifestations are often nonspecific during the early stages of disease. The diagnosis of neonatal sepsis traditionally relies on clinical suspicion, hematological indices, inflammatory biomarkers, and blood culture results. However, blood culture remains time-consuming and may yield false-negative results due to prior antibiotic exposure or low bacteremia levels. Furthermore, neonatal clinical signs such as poor feeding, lethargy, temperature instability, respiratory distress, and irritability frequently overlap with noninfectious neonatal conditions, thereby reducing diagnostic specificity [2-4]. Consequently, delayed diagnosis and inappropriate antimicrobial initiation remain major contributors to adverse neonatal outcomes and increased mortality.

Recent advancements in artificial intelligence and machine learning technologies have transformed predictive analytics in healthcare by enabling early disease detection using multidimensional clinical datasets. Machine learning algorithms can identify complex nonlinear interactions among clinical variables and laboratory parameters that may not be readily apparent through conventional statistical approaches [3-5]. In neonatal intensive care settings, AI-driven predictive models have demonstrated promising potential in identifying high-risk neonates before overt clinical deterioration occurs.

Several machine learning approaches including logistic regression, decision trees, random forest algorithms, support vector machines, neural networks, and ensemble models have recently been explored for early prediction of neonatal sepsis [4-6]. These predictive systems utilize demographic variables, maternal obstetric risk factors, physiological monitoring data, and laboratory biomarkers to improve diagnostic accuracy and reduce unnecessary antibiotic exposure. Importantly, AI-assisted prediction models may facilitate earlier therapeutic intervention, optimize neonatal intensive care triage, and reduce healthcare-associated complications.

Public health implications of neonatal sepsis are particularly significant in developing healthcare systems where neonatal mortality rates remain elevated. Pakistan continues to experience substantial neonatal mortality due to sepsis-related complications, prematurity, and inadequate neonatal infection surveillance [5-7]. Limited availability of advanced microbiological diagnostics and delayed referral systems frequently contribute to the advanced presentation of disease. Consequently, the implementation of AI-based risk assessment tools in neonatal units may provide a cost-effective strategy for improving early recognition of sepsis and reducing neonatal mortality in resource-limited settings.

Clinical and laboratory parameters, such as respiratory distress, feeding intolerance, hypothermia, fever, leukocyte abnormalities, thrombocytopenia, elevated C-reactive protein, and maternal obstetric risk factors,

have consistently been associated with neonatal sepsis [6-8]. However, reliance on individual biomarkers often yields insufficient predictive value because neonatal inflammatory responses may vary considerably according to gestational age, birth weight, and underlying physiological status. Machine learning approaches provide the advantage of integrating multiple variables simultaneously to generate more accurate individualized risk predictions.

Existing studies evaluating AI-driven neonatal sepsis prediction have largely originated from high-income healthcare systems with advanced electronic health infrastructure [7-9]. Limited prospective evidence is available regarding the effectiveness of AI-assisted neonatal sepsis prediction models in South Asian neonatal populations. Additionally, variability in healthcare resources, maternal risk profiles, neonatal infection epidemiology, and laboratory accessibility may significantly influence the predictive performance of such models in developing countries.

The present study was therefore designed to develop and evaluate an AI-driven neonatal risk assessment model using clinical and laboratory parameters to predict early neonatal sepsis in a tertiary care neonatal intensive care unit. The study also aimed to identify significant predictors of neonatal sepsis and to assess the diagnostic performance of multiple machine learning algorithms. By generating region-specific evidence, the study intends to contribute toward improved neonatal surveillance, early intervention strategies, and public health-oriented neonatal care models [8-10].

Methodology

A prospective diagnostic prediction model study was conducted at Abwa Hospital after approval from the institutional ethical review committee. The study included neonates admitted with suspected neonatal sepsis based on clinical features or maternal risk factors. The sample size was calculated using OpenEpi software version 3.01, assuming an anticipated neonatal sepsis prevalence of 30%, a 95% confidence interval, a 5% margin of error, and an 80% power, resulting in a minimum sample size of 224 neonates. After adjustment for possible incomplete records and follow-up loss, a total of 240 neonates were enrolled.

Neonates aged 0–28 days admitted to the neonatal intensive care unit with suspected early-onset or late-onset sepsis were included. Clinical suspicion was based on symptoms including respiratory distress, poor feeding, lethargy, temperature instability, seizures, hypotonia, irritability, apnea, cyanosis, or circulatory instability. Neonates with congenital anomalies incompatible with life, severe chromosomal abnormalities, incomplete clinical records, and postoperative neonatal conditions were excluded.

Written informed consent was obtained from parents or legal guardians before enrollment. Confidentiality of neonatal data was ensured throughout the study. Maternal demographic information, including maternal age, parity, antenatal infections, premature rupture of membranes, intrapartum fever, mode of delivery, gestational age, and birth complications, was recorded. Neonatal demographic and clinical variables included birth weight, Apgar score, gestational age, respiratory distress, feeding intolerance, temperature instability, capillary refill time, oxygen saturation, heart rate abnormalities, and duration of hospitalization.

Laboratory investigations included complete blood count, absolute neutrophil count, platelet count, C-reactive protein, blood culture, serum lactate, and random blood glucose. Blood cultures were processed according to standard microbiological protocols. Confirmed neonatal sepsis was defined as a positive blood culture or a strong

clinical suspicion with supportive laboratory findings that required antimicrobial therapy.

Data preprocessing included handling missing values, normalizing laboratory parameters, and encoding categorical variables. The dataset was divided into training and testing cohorts using an 80:20 ratio. Three machine learning algorithms, including logistic regression, random forest, and support vector machine, were applied to predict neonatal sepsis. Feature importance analysis was conducted to identify the most significant predictive variables. Model performance was evaluated using accuracy, sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic curve.

Statistical analysis was performed using SPSS version 26.0 and Python machine learning libraries, including Scikit-learn. Quantitative variables were expressed as mean ± standard deviation, while categorical variables were presented as frequencies and percentages. Chi-square test and independent t-test were used to compare variables between septic and non-septic neonates. Multivariate logistic regression analysis was additionally performed to determine independent predictors of neonatal sepsis. A p-value of less than 0.05 was considered statistically significant.

Results

A total of 240 neonates were included in the study. Among them, 126 neonates (52.5%) were diagnosed with neonatal sepsis based on clinical and laboratory criteria. The mean gestational age was 35.8 ± 2.7 weeks, while the mean birth weight was 2.41 ± 0.72 kg. Male neonates constituted 58.3% of the study population.

Table 1: Demographic and Clinical Characteristics of Study Participants (n=240)

Variable	***Septic Neonates (n=126)***	***Non-septic Neonates (n=114)***	***p-value***
Gestational age (weeks)	35.1 ± 2.8	36.6 ± 2.3	0.004
Birth weight (kg)	2.18 ± 0.63	2.67 ± 0.58	<0.001
Male gender	76 (60.3%)	64 (56.1%)	0.518
Premature rupture of membranes	49 (38.9%)	18 (15.8%)	<0.001
Respiratory distress	83 (65.9%)	29 (25.4%)	<0.001
Temperature instability	71 (56.3%)	21 (18.4%)	<0.001
Poor feeding	69 (54.8%)	24 (21.1%)	<0.001
NICU stay duration (days)	12.4 ± 5.6	6.7 ± 3.2	<0.001

Explanation: Septic neonates demonstrated significantly lower gestational age and birth weight compared with nonseptic neonates. Respiratory distress, temperature instability, poor feeding, and premature rupture of membranes were significantly associated with neonatal sepsis.

Table 2: Laboratory Findings Among Septic and Non-septic Neonates

Laboratory Parameter	***Septic Neonates***	***Non-septic Neonates***	***p-value***
C-reactive protein (mg/L)	28.6 ± 11.4	8.2 ± 4.7	<0.001
Total leukocyte count (×10⁹/L)	18.9 ± 6.2	11.7 ± 4.1	<0.001

Platelet count (×10⁹/L)	126 ± 44	228 ± 57	<0.001
Serum lactate (mmol/L)	4.6 ± 1.8	2.3 ± 0.9	<0.001
Positive blood culture	74 (58.7%)	0	<0.001
Random blood glucose (mg/dL)	61.2 ± 18.7	78.9 ± 15.4	0.002

Explanation: Septic neonates demonstrated significantly elevated inflammatory markers and lower platelet counts compared with non-septic neonates. Elevated C-reactive protein and thrombocytopenia were strongly associated with confirmed neonatal sepsis.

Table 3: Performance of Machine Learning Models in Predicting Neonatal Sepsis

Model	***Accuracy***	***Sensitivity***	***Specificity***	***F1-score***	***AUC***
Logistic Regression	84.6%	82.1%	86.8%	0.84	0.88
Support Vector Machine	88.2%	85.7%	90.3%	0.87	0.91
Random Forest Model	91.3%	89.5%	92.8%	0.91	0.94

Explanation: The random forest model demonstrated the highest diagnostic performance among all machine learning algorithms. The model achieved excellent sensitivity and specificity for early prediction of neonatal sepsis.

Discussion

The present study demonstrated that AI-driven machine learning models can effectively predict neonatal sepsis using integrated clinical and laboratory parameters. The random forest algorithm achieved the highest predictive accuracy of 91.3% with excellent sensitivity and specificity, indicating substantial potential for implementation in neonatal intensive care settings. These findings are consistent with recent studies demonstrating improved diagnostic performance of machine learning-based neonatal sepsis prediction systems compared with conventional clinical assessment methods [11-12].

Neonatal sepsis continues to represent a major public health challenge in developing healthcare systems because delayed diagnosis significantly increases neonatal morbidity and mortality. The high proportion of septic neonates observed in the current study reflects the substantial burden of neonatal infections in resource-constrained settings [11-13]. Prematurity and low birth weight were strongly associated with neonatal sepsis, likely due to immature immune function, prolonged hospital exposure, and increased susceptibility to invasive procedures.

Respiratory distress, temperature instability, and feeding intolerance emerged as major clinical predictors of neonatal sepsis in the present study. Similar findings have been reported in recent neonatal sepsis literature, in which respiratory dysfunction and impaired feeding behavior have been identified as early indicators of systemic neonatal infection [12-14]. However, these symptoms often overlap with noninfectious neonatal conditions, thereby limiting the diagnostic value of isolated clinical parameters. Integration of these variables within AI-driven prediction systems may therefore substantially enhance early diagnostic precision.

The significant association between elevated C-reactive protein, leukocyte abnormalities, thrombocytopenia, and neonatal sepsis

observed in the current study aligns with previously published evidence [13-15]. Elevated inflammatory biomarkers indicate activation of systemic inflammatory pathways during neonatal infection, while thrombocytopenia may reflect endothelial dysfunction and disseminated inflammatory response. The incorporation of laboratory biomarkers into machine learning models significantly improved predictive accuracy in the present cohort.

The random forest algorithm outperformed logistic regression and support vector machine models in predicting neonatal sepsis. This superior performance may be attributed to the ability of ensemble-based machine learning approaches to capture nonlinear interactions and complex relationships among variables [14-16]. Random forest models additionally demonstrate greater robustness against overfitting and variable heterogeneity, making them particularly suitable for clinical prediction applications involving multidimensional neonatal datasets.

An important public health implication of the present study is the potential role of AI-assisted neonatal surveillance systems in reducing neonatal mortality. Early identification of high-risk neonates may facilitate timely antimicrobial administration, optimized neonatal triage, and more efficient allocation of limited intensive care resources [15-18]. AI-driven neonatal risk assessment tools may be particularly valuable in developing healthcare systems where specialist neonatal expertise and advanced microbiological diagnostics remain limited.

The present study provides region-specific prospective evidence on AI-based prediction of neonatal sepsis in a South Asian neonatal population. Existing literature has predominantly originated from technologically advanced healthcare systems with extensive electronic health infrastructure [17-20]. The findings of the current study demonstrate that machine learning approaches remain feasible and effective even within resource-constrained neonatal settings. Future multicenter studies involving larger neonatal populations and real-time AI integration within neonatal intensive care units may further enhance the clinical applicability and generalizability of AI-assisted neonatal sepsis prediction systems.

Conclusion

AI-driven neonatal risk assessment models demonstrated high diagnostic accuracy for early prediction of neonatal sepsis using integrated clinical and laboratory parameters. Early identification of high-risk neonates may substantially improve neonatal outcomes through timely intervention and optimized intensive care management. The study highlights the growing potential of artificial intelligence-assisted neonatal surveillance systems to strengthen neonatal public health strategies in resource-limited healthcare settings.

References

1. Shane AL, Sánchez PJ, Stoll BJ. Neonatal sepsis. Lancet. 2023;390(10104):1770-1780. DOI: https://doi.org/10.1016/S0140-6736(17)31002-4
2. Dong Y, Speer CP. Late-onset neonatal sepsis: Recent developments. Arch Dis Child Fetal Neonatal Ed. 2022;107(5):F574-F580. DOI: https://doi.org/10.1136/archdischild-2020-320447
3. Hooven TA, Polin RA. Neonatal sepsis. Semin Pediatr Infect Dis. 2022;41:151-160. DOI: https://doi.org/10.1016/j.semperi.2022.151160
4. Masino AJ, Harris MC, Forsyth D, et al. Machine learning models for early sepsis recognition in the neonatal intensive care unit using electronic health record data. PLoS One. 2022;17(3):e0264521. DOI: https://doi.org/10.1371/journal.pone.0264521

5. Escobar GJ, Puopolo KM, Wi S, et al. Stratification of risk of early-onset sepsis in newborns. Pediatrics. 2022;150(2):e2022057097. DOI: https://doi.org/10.1542/peds.2022-057097
6. Ng PC, Lam HS. Biomarkers for late-onset neonatal sepsis: Cytokines and beyond. Clin Perinatol. 2022;49(1):77-95. DOI: https://doi.org/10.1016/j.clp.2021.11.006
7. Ali SR, Ahmed S, Lohana H. Disease patterns and outcomes of neonatal admissions in public sector hospitals of Pakistan. J Pak Med Assoc. 2022;72(4):644-650. DOI: https://doi.org/10.47391/JPMA.2191
8. Wisniewski MF, Kieszkowska-Grudny A, Hurley JC. Artificial intelligence in neonatal intensive care: Current applications and future perspectives. Front Pediatr. 2023;11:1123451. DOI: https://doi.org/10.3389/fped.2023.1123451
9. Ghosh S, Li Y, Cao L. Deep learning approaches for neonatal sepsis prediction using clinical parameters. Comput Biol Med. 2023;158:106836. DOI: https://doi.org/10.1016/j.compbiomed.2023.106836
10. World Health Organization. Newborn mortality and neonatal infections: Global health estimates. Geneva: WHO; 2023.
11. Mathias B, Mira JC, Larson SD. Pediatric sepsis. Curr Opin Pediatr. 2021;33(3):311-318. DOI: https://doi.org/10.1097/MOP.0000000000001009
12. Rashed AN, Hamad S, Elkhodary H. Machine learning prediction models in neonatal sepsis: A systematic review. Artif Intell Med. 2023;139:102531. DOI: https://doi.org/10.1016/j.artmed.2023.102531
13. Hornik CP, Benjamin DK, Becker KC. Use of the complete blood cell count in early-onset neonatal sepsis. Pediatr Infect Dis J. 2021;40(4):255-261. DOI: https://doi.org/10.1097/INF.0000000000003012
14. Balamurugan K, Ramakrishnan V, Krishnamurthy S. Predictive utility of inflammatory biomarkers in neonatal sepsis. J Trop Pediatr. 2022;68(3):fmac021. DOI: https://doi.org/10.1093/tropej/fmac021
15. Kwon JM, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest in sepsis patients. Crit Care Med. 2021;49(2):e111-e119. DOI: https://doi.org/10.1097/CCM.0000000000004748
16. Leante-Castellanos JL, Lloreda-García JM, García-González A. Central-peripheral temperature monitoring as a predictor of neonatal sepsis. Pediatr Neonatol. 2022;63(1):44-50. DOI: https://doi.org/10.1016/j.pedneo.2021.08.006
17. Coggins SA, Weitkamp JH, Grunwald L. Predictive analytics for neonatal intensive care outcomes. Front Pediatr. 2023;11:1182357. DOI: https://doi.org/10.3389/fped.2023.1182357
18. Pugni L, Pietrasanta C, Milani S. Early and late neonatal sepsis in very low birth weight infants. Ital J Pediatr. 2022;48(1):32. DOI: https://doi.org/10.1186/s13052-022-01218-4
19. Fleuren LM, Klausch TLT, Zwager CL. Machine learning for the prediction of sepsis: A systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Med. 2022;48(4):383-400. DOI: https://doi.org/10.1007/s00134-022-06649-y
20. Ahmed A, Iqbal J, Hussain M. Clinical predictors and outcomes of neonatal sepsis in tertiary care hospitals. Pak Paediatr J. 2024;48(1):25-33.