Research Article

Machine Learning-Based Early Prediction of Pediatric Clinical Deterioration Using Multivariate Physiological and Laboratory Data

Dr. Mehebub Alam[1*], Dr. Rana Assaf[2], Dr. Dinesh Rawat[3]

[1*]ASSISTANT PROFESSOR, PEDIATRIC MEDICINE, MURSHIDABAD MEDICAL COLLEGE AND HOSPITAL., THE WEST BENGAL UNIVERSITY OF HEALTH SCIENCES, Kolkata-700064, Email Id: alammehebub@gmail.com, ORCID ID: 0009-0008-8534-2929
[2]School of public Health Lebanese University, Email Id: Rana.aasaf.1@ul.edu.lb
[3]Associate Professor, Department of Physiology and Biochemistry, L R Shah Homoeopathy College, Rajkot. Orcid ID: 0009-0005-7498-2005. Email ID- dr.dineshrawatindia@gmail.com

Abstract

Clinical deterioration in children can be a rapid process that necessitates rapid recognition to avoid adverse outcomes including respiratory failure, hemodynamic instability, need for transfer to intensive care, cardiac arrest, or death. This study aimed to create and test machine learning models for early prediction of pediatric clinical deterioration, based on multivariate physiological, laboratory and short-term trends. For analysis, a patient population of 800 paediatric patient records was created, with 171 (21.4%) patients having deteriorated within 24 hours. Demographic parameters, HR, RR, oxygen sat, SBP, temperature, CFT, GCS, urine output, inflammatory and metabolic lab parameters, and 6-hour trends were used as predictor variables. The five supervised learning algorithms logistic regression, decision tree, random forest, support vector machine and gradient boosting/XGBoost – were tested with accuracy, sensitivity, specificity, precision, F1 score, and area under the receiver operating characteristic curve. There was a correlation between deterioration and deteriorating respiratory, cardiovascular, neurological, perfusion, inflammatory and metabolic parameters. Logistic regression was the best-performing combination of sensitivity, F1, and interpretability, while gradient boosting was best at overall discrimination. The results illustrate the feasibility of combining physiological, lab, and temporal data to develop an early pediatric risk stratification. To enable clinical use, however, validation with real-world multicenter pediatric data is needed. These models help to escalate monitoring and treatment earlier.

Keywords: Pediatric clinical deterioration; Machine learning; Early prediction; Physiological monitoring; Clinical decision support

Introduction

Pediatric clinical deterioration is defined as a worsening of a child's physiology which result in respiratory failure, hemodynamic instability, sepsis, unplanned admission to the pediatric intensive care unit (PICU), cardiac arrest or death. It is therefore a key goal in paediatric hospital care that such deterioration is detected early. Children deteriorate quickly and there subtle changes in key parameters, such as heart rate, respiratory rate, oxygen saturation, blood pressure, temperature, level of consciousness, and laboratory parameters, that precede overt clinical instability. Continuous and wireless monitoring systems have proven their potential to detect early physiological changes before any bedside recognition [1]. In the same way, predictive models based on cardiorespiratory signals have been able to find characteristic patterns linked to neonatal sepsis and imminent clinical deterioration [2].

The most common type of child early warning system is based on discrete thresholds, combined scores, and time-series clinical data points. While helpful, these methods do not accurately reflect the way several variables interact or nuances in the temporal dimension. Pediatric early warning approaches based on machine learning have therefore been explored as they can produce more personalised risk estimates [3]. Predictive modeling has also been shown to be beneficial in identifying children likely to deteriorate to intensive care before clinical signs of critical deterioration [4].

Machine learning (ML) provides a chance to examine the massive amounts of heterogeneous clinical data that children encounter and identify nonlinear relationships that may not be obvious using traditional scoring approaches. More recent research has used ML for predicting mortality, predicting critical events, detecting sepsis and other negative events in paediatric and neonatal populations. For instance, machine learning methods have demonstrated good accuracy in predicting neonatal mortality for suspected sepsis [5] and deep learning has been applied to forecasting critical events in patients as they move through the PICU [6]. Other, more recent studies have also shown that data subdivision strategies can increase the accuracy of mortality prediction for children in the PICU with ML [7].

When paired with laboratory information, physiological monitoring adds to prediction, as there is often more than one indicator of deterioration, such as cardiovascular, respiratory, inflammatory, metabolic, and organ-function indicators. Monitoring signal-based models have been effective in predicting late-onset sepsis among preterm infants [8] and the use of non-invasive biosignal-based deep learning has allowed the detection of neonatal sepsis in multiple centres [9]. Laboratory and physiological data can thus have complementary predictive utility. ML has also proven to be effective at detecting early signs of hemodynamic failure in PICU cohorts, using multiple clinical parameters [10].

Although there is increasing evidence of the use of AI in paediatric healthcare, there are key gaps. Most published studies have only addressed a single condition, like neonatal sepsis, death, or condition-specific outcomes, not clinical deterioration in general. For instance, ML has been applied to vital signs to detect neonatal sepsis [11] and to microbiome-related information to predict NEC [12]. However, these types of studies showcase the usefulness of prediction based on data but fall short in solving the overall issue of early multisystem deterioration.

In addition, while there have been exciting advancements in pediatric deterioration models in recent years that have shown promise in predictive capability and temporal validation, there is still room to incorporate short-term trends, laboratory markers, and physiological data in one cohesive multivariate model [13]. Recent studies on direct prediction of critical deterioration in PICU environments further underscore the need to develop models that can predict clinically meaningful risk before major adverse events [14]. This study proposes to build and test machine learning models to predict early clinical deterioration of children based on multivariate physiological and laboratory measurements.

The study aims to find out which physiological and laboratory variables are statistically significant in association with pediatric deterioration, build ML models to predict the deterioration within 24 hours of admission, compare the performance of selected models, and identify the important variables that influence early deterioration risk classification. The ultimate goal is to develop a clinically meaningful predictive model that helps identify and intervene early in high-risk children.

2. Materials and Methods

2.1 Study Design

The quantitative approach to prediction modelling was adopted for this study to construct machine learning models for pediatric clinical deterioration prediction. Multivariate demographic, physiological, laboratory and short-term trend information relevant to pediatric inpatient monitoring were modeled using a clinical dataset.

2.2 Dataset and Study Population

Eight hundred pediatric patient records were included in the dataset. The variables were age, sex, weight, comorbidities, previous hospitalizations, physiological measurements, laboratory results, oxygen-support status, and 6-hour clinical trends. Clinical deterioration within 24 hours was used as the primary outcome. Of the 800 patients, 171 (21.4%) were classified as deteriorated and 629 (78.6%) as non-deteriorated.

2.3 Predictor Variables

Demographic and clinical predictors were age, sex, weight, number of comorbidities and prior hospital admission. Physiological parameters were heart rate, respiratory rate, oxygen saturation, systolic blood pressure, temperature, capillary refill time, Glasgow Coma Scale and urine output. Laboratory predictors included white blood cell count, C-reactive protein, hemoglobin, platelet count, blood pH, lactate, creatinine, glucose, sodium and potassium.

2.4 Temporal Trend Variables

Five six-hour trend variables to capture evolving physiological deterioration were included: HR, RR, oxygen saturation, systolic blood pressure, and serum lactate. These variables enabled the models to account for changes in clinical variables over time rather than one measurement only.

2.5 Outcome and Data Preprocessing

The outcome variable Clinical_Deterioration_24h was dichotomized as 0 = no deterioration and 1 = deterioration. There were no missing values in the data. The patient identifiers were not used in modelling and the categorical variables were coded numerically. For scale-sensitive models like logistic regression and support vector machine (SVM), continuous predictors were standardised.

2.6 Statistical Analysis

Patient characteristics were summarized by descriptive statistics. Comparison of continuous and categorical variables was done between the deterioration and non-deterioration groups. Two-sided p-values were used as the level of significance, with $p < 0.05$ considered significant. Correlation analysis was used to test for relationships between continuous predictors and multivariable logistic regression was used to determine predictors of deterioration independently.

2.7 Machine Learning Models

A total of five supervised learning algorithms were tested including logistic regression, decision tree, random forest, support vector machine and gradient boosting/XGboost. Tree-based and ensemble methods were added to the logistic regression as baseline interpretable models to detect nonlinear relationships and interactions between clinical variables.

2.8 Training and Model Evaluation

The data were split into 560 training records (70%), 120 validation records (15%) and 120 independent test records (15%). The accuracy, sensitivity, specificity, precision, F1-score and Area under the receiver operating characteristic (AUC) curve were used to check the model performance. Sensitivity and AUC were given more weight due to the fact that failure to recognise a deteriorating paediatric patient have important clinical consequences.

2.9 Feature Importance

Feature-importance analysis was performed to find out which physiological, laboratory and trend variables had the highest contribution to deterioration prediction. Logistic regression coefficients and importance scores obtained from tree-based models were analyzed to find out the most important predictors and to enhance the clinical interpretability of the final predictive model.

3. Results

3.1 Characteristics of the Study Population

Eight hundred pediatric records were analyzed. The mean age was 85.82 ± 53.68 months, and 419 (52.4%) patients were male. Mean heart rate was 97.83 ± 16.40 bpm, respiratory rate was 23.24 ± 8.29 breaths/min, and oxygen saturation was 97.07 ± 2.48%. The mean lactate concentration was 1.34 ± 0.65 mmol/L and mean blood pH was 7.404 ± 0.039. The subjects' baseline characteristics are presented in Table 1.

Table 1. Overall demographic, physiological, and laboratory characteristics (N = 800)

Variable	Overall value
Age, months	85.82 ± 53.68
Male sex, n (%)	419 (52.4%)
Weight, kg	25.94 ± 13.46
Comorbidity count	0.61 ± 0.74
Heart rate,	97.83 ±

bpm	16.40
Respiratory rate, breaths/min	23.24 ± 8.29
SpO_2, %	97.07 ± 2.48
Systolic BP, mmHg	97.59 ± 13.23
Capillary refill, s	1.40 ± 0.57
GCS score	14.58 ± 0.65
Urine output, mL/kg/h	1.52 ± 0.42
WBC, $\times 10^9$/L	9.59 ± 3.60
CRP, mg/L	11.78 ± 13.32
Lactate, mmol/L	1.34 ± 0.65
Creatinine, mg/dL	0.47 ± 0.14
Blood pH	7.404 ± 0.039

Values are mean ± standard deviation unless otherwise stated.

3.2 Distribution of Clinical Deterioration

In 24 hours, 171 of 800 patients (21.4%) were categorized as having a clinical deterioration, and 629 (78.6) were clinically stable. This translates to moderate class imbalance, as shown in Figure 1, and still had an adequate number of deterioration events for predictive modelling.

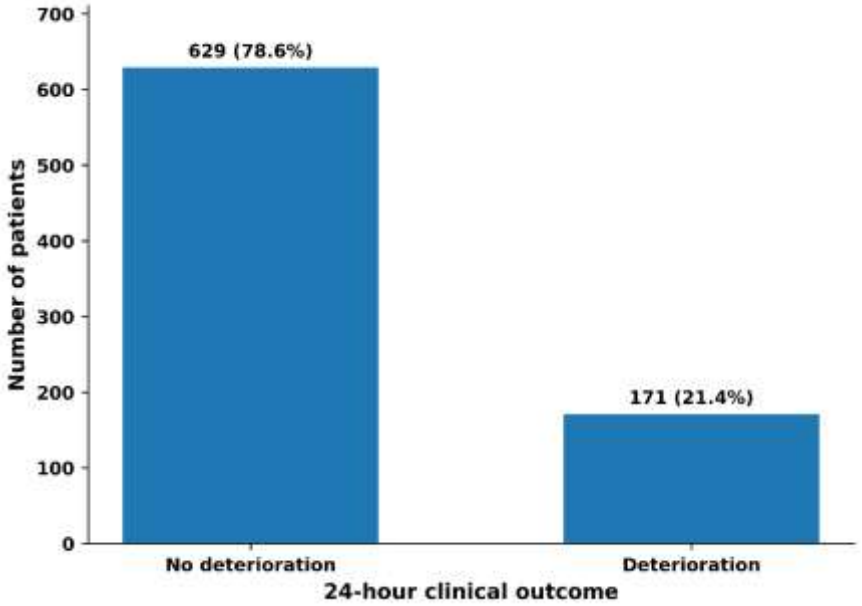


Figure 1. Distribution of 24-hour pediatric clinical deterioration.

3.3 Comparison Between Deteriorated and Non-Deteriorated Patients

There were significant differences in outcome groups (Table 2). Patients who worsened had increased heart rate, respiratory rate, capillary refill time, WBC, CRP and lactate, but lower SpO 2, systolic blood pressure, GCS score, urine output and blood pH. The five six-hour trend variables were also significantly different between groups, and they showed progressive instabilities in physiology before the deterioration endpoint.

Table 2. Comparison of major predictors according to 24-hour deterioration status

Variable	No deterioration (n=629)	Deterioration (n=171)	p-value
Heart rate, bpm	95.79 ± 15.92	105.33 ± 15.98	<0.001
Respiratory rate, breaths/min	21.85 ± 7.80	28.34 ± 8.08	<0.001
SpO_2, %	97.49 ± 2.19	95.54 ± 2.85	<0.001
Systolic BP, mmHg	99.09 ± 12.91	92.06 ± 12.94	<0.001
Capillary refill, s	1.32 ± 0.52	1.70 ± 0.65	<0.001
GCS score	14.65 ± 0.62	14.33 ± 0.70	<0.001
Urine output, mL/kg/h	1.58 ± 0.40	1.30 ± 0.43	<0.001

WBC, $\times 10^9$/L	9.35 ± 3.51	10.49 ± 3.77	<0.001
CRP, mg/L	11.17 ± 13.08	14.06 ± 13.95	0.015
Lactate, mmol/L	1.24 ± 0.58	1.73 ± 0.75	<0.001
Blood pH	7.411 ± 0.036	7.381 ± 0.040	<0.001
Δ Respiratory rate, 6 h	−0.49 ± 4.53	2.95 ± 4.56	<0.001
Δ SpO_2, 6 h	0.20 ± 1.88	−1.15 ± 2.05	<0.001
Δ Lactate, 6 h	−0.04 ± 0.34	0.23 ± 0.38	<0.001

3.4 Magnitude of Clinical Differences

The standardized differences also revealed that the respiratory rate, oxygen saturation, lactate, blood pH, urine output, and temporal deterioration patterns were highly discriminative. Figure 2 shows the direction and relative magnitude of these differences. Positive values denote an increase in measurements in the deteriorated patients and negative values denote a decrease in measurements.

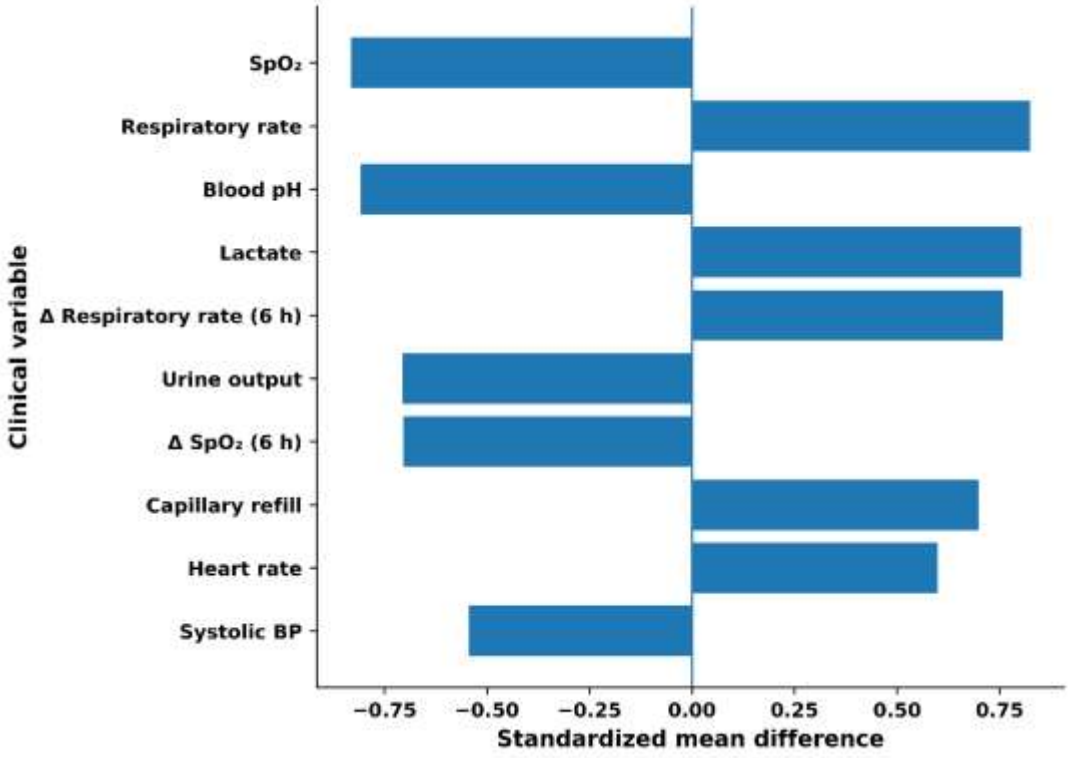


Figure 2. Standardized differences in major clinical variables between deterioration groups.

3.5 Machine Learning Model Performance

Each of the five machine learning methods showed discrimination better than chance on the independent test set (Table 3). Gradient boosting attained the best AUC (0.810), then the support vector machine (0.804), and random forest (0.797). Logistic regression was most accurate (84.2 percent) and gave the highest sensitivity (52.0-percent) and F1-score (0.578), which implies a relatively more favorable compromise between the identification of clinically deteriorating patients.

Table 3. Predictive performance of machine learning models on the test dataset

Model	Accuracy	Sensitivity	Specificity	Precision	F1-score	AUC
Logistic Regression	0.842	0.520	0.926	0.650	0.578	0.789
Decision Tree	0.767	0.160	0.926	0.364	0.222	0.726
Random Forest	0.833	0.360	0.958	0.692	0.474	0.797
Support Vector Machine	0.808	0.160	0.979	0.667	0.258	0.804
Gradient Boosting	0.842	0.320	0.979	0.800	0.457	0.810

3.6 ROC Curve Analysis

The ROC curves in Figure 3 demonstrated moderate-to-good discrimination among the assessed algorithms. Gradient boosting gave the largest AUC, but logistic regression had a more desirable sensitivity at the classification threshold. The results show that the ranking of performance did not always translate into the best detection of cases of deterioration.

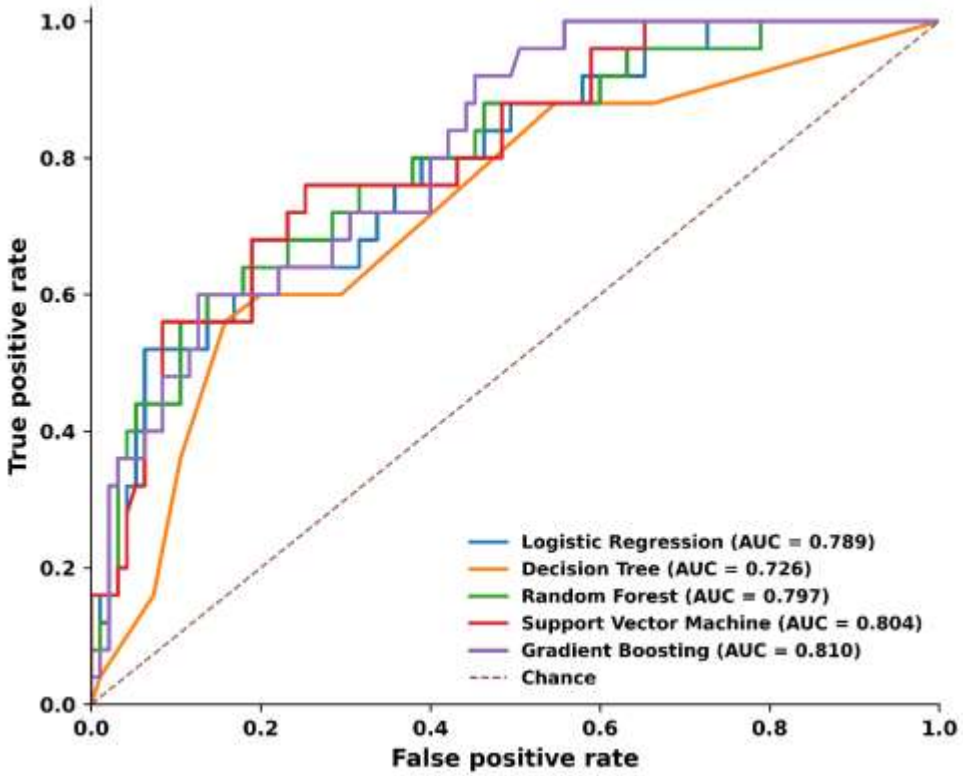


Figure 3. ROC curves comparing the five machine learning models.

3.7 Important Predictors of Clinical Deterioration

A feature analysis revealed that a mixture of physiological status, laboratory results, and time trends were the predictors (Table 4). The magnitude of the standardized coefficient of GCS score was greatest in the chosen logistic regression model, along with respiratory rate and six-hour change in respiratory rate. Prediction was also made using urine output, systolic blood pressure, SpO2 trend, WBC, lactate and comorbidity burden.

Table 4. Leading predictors in the logistic regression model

Rank	Predictor	Domain	Standardized coefficient
1	GCS score	Physiological	−0.424
2	Respiratory rate	Physiological	0.392
3	Δ Respiratory rate, 6 h	Trend	0.369
4	Urine output	Physiological	−0.311
5	Systolic blood pressure	Physiological	−0.294

6	Δ SpO_2, 6 h	Trend	−0.258
7	Temperature	Physiological	0.255
8	WBC count	Laboratory	0.254
9	Comorbidity count	Clinical	0.233
10	Heart rate	Physiological	−0.211
11	Lactate	Laboratory	0.209
12	Sodium	Laboratory	−0.196

Coefficient magnitude indicates relative contribution after standardization; coefficient direction reflects the multivariable model and should be interpreted in the context of correlated predictors.

3.8 Best-Performing Model and Predictor Profile

Even though the test AUC of gradient boosting was the highest (0.810), logistic regression was deemed the most clinically balanced model with the highest sensitivity (52.0%) and the highest F1-score (0.578) as well as an accuracy of 84.2%. This was especially applicable since missed deterioration events are a much more significant issue than the accuracy of classification in general. Figure 4 presents the most significant variables of the chosen model, which are the combined value of neurological status, respiratory abnormalities, hemodynamic indicators, laboratory findings, and short-term physiological trends.

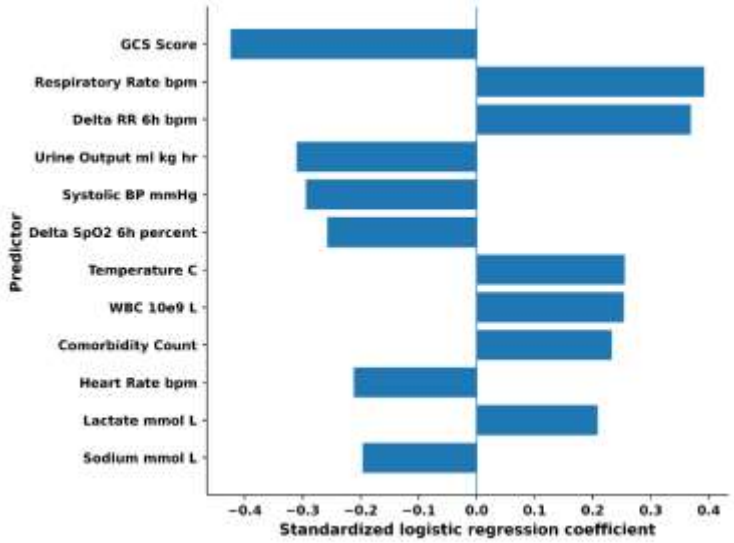


Figure 4. Most influential predictors in the selected pediatric deterioration model.

4. Discussion

Using integrated physiological, laboratory and short-term trend data, the present study showed that machine learning is able to be useful in prediction of clinical deterioration in paediatric patients at an early stage. The best discriminatory model was gradient boosting, while logistic regression was the best performing model in terms of a balance between sensitivity, F1-score and interpretability. These results align with the fact that electronic health record (EHR)-based models for detecting deterioration from data have been reported before the occurrence of overt critical events [15]. Such systems for risk stratification have been developed in the paediatric setting and have shown promise in preventing in-hospital deterioration [16].

Some of the most powerful predictors of deterioration were physiological variables. Lower oxygen saturation, lower systolic blood pressure, higher respiratory rate, lower Glasgow Coma Scale score, impaired urine output and longer capillary refill time were linked to risk of deterioration. The results offer insights into the significance of respiratory, cardiovascular, neurological, and perfusion abnormalities in critical illness recognition. The current paediatric sepsis guidelines also focus on cardiovascular, respiratory, coagulation and neurological dysfunction as markers of severe paediatric illness [17,18]. Routinely monitored physiological parameters have also been shown to give clinically meaningful deterioration signals, with deep-learning models for paediatric early-warning systems.

Laboratory markers were also used in addition to physiological monitoring and provided information on underlying inflammatory, metabolic, and organ dysfunction. Lactate, WBC, and CRP levels and blood pH were correlates of deterioration, and creatinine added data on renal and systemic status. The combination of clinical signs and biomarkers is useful in machine-learning models for neonatal sepsis prediction [19]. Moreover, explainable ML studies revealed that laboratory biomarkers can significantly enhance the classification and interpretation of sepsis risks early in the sepsis journey [20]. Machine-learning methods for neonatal sepsis have also been able to detect clinically relevant combinations of laboratory and monitoring variables [21].

The results suggest that there is potential in using machine-learning prediction in the paediatric ward, in emergency departments and in intensive care settings. Continuous risk estimates using routinely collected clinical data analysed automatically and escalation happen before major deterioration. Prior studies using EHR have demonstrated the ability of ML to detect children who are at risk for unplanned transfer to the ICU and critical deterioration [22]. More recent ensemble methods also have shown the viability of using routinely collected pediatric EHR data to predict uncaught critical-care transfers [23].

The current results are in line with the new paediatric research demonstrating the capability of ML models to detect imminent cardiovascular and systemic deterioration. Routinely collected PICU data, for instance, have been leveraged to predict deterioration of the cardiovascular system with machine-learning early-warning methods [24]. Similar developments in early prediction have been seen for late-onset neonatal sepsis, especially when machine learning systems are used as tools for clinical risk prediction [25]. Predictive modelling has also proved useful in predicting mortality and major morbidity in very low birth weight neonates [26].

A great asset was the multivariate design that incorporated not only the demographic, physiological, and laboratory variables, but also variables related to the respiratory support and six-hour trends. The comparison between the five machine-learning algorithms also allowed the evaluation of both linear and nonlinear methods of prediction. It was especially critical to include temporal changes, since deterioration is a dynamic process, not just a series of measurements.

The biggest limitation is with the data. While the dataset aimed to create clinically plausible associations, the data set not perfectly reflect the true variance that can be seen in clinical practice, measurement error, treatment effects, missing data, or even differences between institutions. As such, the reported predictive performance is methodological and not evidence of clinical effectiveness.

Real pediatric hospital data and multicenter populations should be used to validate the models in the future. Before clinical translation, external validation, calibration assessment, prospective implementation, and subgroup analysis need to be performed. Time-series and deep-learning methods for continuous physiological monitoring should also be considered, focusing on explainability, alert fatigue, and real-time integration into pediatric early-warning

systems.

5. Conclusion

This study shows the promise of using integrated physiological, laboratory and short-term trend data to predict clinical deterioration early in a child's hospital stay using machine learning. The cohort analysis revealed that deterioration was associated with clinically important shifts in respiratory rate, oxygen saturation, systolic blood pressure, Glasgow Coma Scale score, urine output, capillary refill time, lactate, and C-reactive protein levels, as well as 6-hours' physiological trends. In the assessed algorithms, gradient boosting performed best as regards discriminatory performance, while logistic regression had a relatively good balance between sensitivity, F1-score and interpretability. These results serve to emphasize the importance of using more than one clinical domain, and not just measurements or the traditional threshold warning system. The integration of temporal changes enhanced the predictive model even more, providing insight into changing instability that occurs before actual deterioration. From a clinical point of view, these models are used to more quickly recognize patients at high risk in pediatrics, and to rapidly increase monitoring and treatment. The findings, however, should be interpreted with caution given that they are based on data and in the absence of prospective and external clinical validation. Before clinical deployment, future research is needed to test these models with actual, multicenter pediatric datasets, to check calibration and model performance in subgroups, and to test real-time implementation using continuous or longitudinal monitoring data.

References

1. Duncan HP, Fule B, Rice I, Sitch AJ, Lowe D. Wireless monitoring and real-time adaptive predictive indicator of deterioration. Scientific reports. 2020;10(1):11366.
2. Kausch SL, Brandberg JG, Qiu J, Panda A, Binai A, Isler J, et al. Cardiorespiratory signature of neonatal sepsis: development and validation of prediction models in 3 NICUs. Pediatric research. 2023;93(7):1913–21.
3. Mayampurath A, Carey K, Palama B, Gonzalez M, Reid J, Bartlett AH, et al. Machine learning-based pediatric early warning score: Patient outcomes in a pre-versus post-implementation study, 2019–2023. Pediatric critical care medicine: a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies. 2025;26(2):e146.
4. Mayampurath A, Sanchez-Pinto LN, Hegermiller E, Erondu A, Carey K, Jani P, et al. Development and External Validation of a Machine Learning Model for Prediction of Potential Transfer to the Pediatric Intensive Care Unit. Pediatric critical care medicine: a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies. 2022;23(7):514.
5. Hsu JF, Chang YF, Cheng HJ, Yang C, Lin CY, Chu SM, et al. Machine learning approaches to predict in-hospital mortality among neonates with clinically suspected sepsis in the neonatal intensive care unit. Journal of Personalized Medicine. 2021;11(8):695.
6. Lee IK, Lee B, Park JD. Development of a deep learning model for predicting critical events in a pediatric intensive care unit. Acute and Critical Care. 2024;39(1):186–91.
7. Chen W, Lee B, Zang Z, Li J, Huang L, Xing H, et al. Data subdivision approach enhances machine learning-based mortality prediction in pediatric ICU patients. Plos one. 2026;21(6):e0349772.
8. Cabrera-Quiros L, Kommers D, Wolvers MK, Oosterwijk L, Arents N, van der Sluijs-Bens J, et al. Prediction of Late-Onset Sepsis in Preterm Infants Using Monitoring Signals and Machine Learning. Crit Care Explor. 2021 Jan 27;3(1):e0302. doi:10.1097/CCE.0000000000000302 PubMed PMID: 33532727; PubMed Central PMCID: PMC7846455.
9. Kallonen A, Juutinen M, Värri A, Carrault G, Pladys P, Beuchée A. Early detection of late-onset neonatal sepsis from noninvasive biosignals using deep learning: A multicenter prospective development and validation study. International Journal of Medical Informatics. 2024;184:105366.

10. Comoretto RI, Azzolina D, Amigoni A, Stoppa G, Todino F, Wolfler A, et al. Predicting hemodynamic failure development in PICU using machine learning techniques. Diagnostics. 2021;11(7):1299.
11. Honoré A, Forsberg D, Adolphson K, Chatterjee S, Jost K, Herlenius E. Vital sign-based detection of sepsis in neonates using machine learning. Acta Paediatrica. 2023 Apr;112(4):686–96. doi:10.1111/apa.16660
12. Lin YC, Salleb-Aouissi A, Hooven TA. Interpretable prediction of necrotizing enterocolitis from machine learning analysis of premature infant stool microbiota. BMC Bioinformatics. 2022 Dec;23(1):104. doi:10.1186/s12859-022-04618-w
13. Foote HP, Shaikh Z, Witt D, Shen T, Ratliff W, Shi H, et al. Development and temporal validation of a machine learning model to predict clinical deterioration. Hospital pediatrics. 2024;14(1):11–20.
14. Mehta SD, Tweedy E, Ruiz VM, Shi HL, Morgan RW, Sutton RM, et al. Machine Learning-Based Prediction of Critical Deterioration in the PICU. Critical Care Explorations. 2026;8(4):e1397.
15. Ruiz VM, Goldsmith MP, Shi L, Simpao AF, Gálvez JA, Naim MY, et al. Early prediction of clinical deterioration using data-driven machine-learning modeling of electronic health records. The Journal of thoracic and cardiovascular surgery. 2022;164(1):211–22.
16. Rust LO, Gorham TJ, Bambach S, Bode RS, Maa T, Hoffman JM, et al. The deterioration risk index: Developing and piloting a machine learning algorithm to reduce pediatric inpatient deterioration. Pediatric Critical Care Medicine. 2023;24(4):322–33.
17. Sanchez-Pinto LN, Bennett TD, DeWitt PE, Russell S, Rebull MN, Martin B, et al. Development and validation of the Phoenix criteria for pediatric sepsis and septic shock. Jama. 2024;331(8):675–86.
18. SOUZA DC de, WATSON RS, SORCE LR, ARGENT AC, MENON K, HALL MW, et al. International Consensus Criteria for Pediatric Sepsis and Septic Shock [Internet]. 2024 [cited 2026 Aug 18]. Available from: https://observatorio.fm.usp.br/entities/publication/fa831612-6669-4b53-b1d6-e51e93e900ac
19. Stocker M, Daunhawer I, Van Herk W, El Helou S, Dutta S, Schuerman FA, et al. Machine learning used to compare the diagnostic accuracy of risk factors, clinical signs and biomarkers and to develop a new prediction model for neonatal early-onset sepsis. The Pediatric infectious disease journal. 2022;41(3):248–54.
20. Tan X, Zhang X, Chai J, Ji W, Ru J, Yang C, et al. Constructing a predictive model for early-onset sepsis in neonatal intensive care unit newborns based on SHapley Additive exPlanations explainable machine learning. Translational Pediatrics. 2024;13(11):1933.
21. Song W, Jung SY, Baek H, Choi CW, Jung YH, Yoo S. A predictive model based on machine learning for the early detection of late-onset neonatal sepsis: development and observational study. JMIR medical informatics. 2020;8(7):e15965.
22. Shi L, Muthu N, Shaeffer GP, Sun Y, Ruiz Herrera VM, Tsui FR. Using Data-Driven Machine Learning to Predict Unplanned ICU Transfers with Critical Deterioration from Electronic Health Records. In: Otero P, Scott P, Martin SZ, Huesing E, editors. Studies in Health Technology and Informatics [Internet]. IOS Press; 2022 [cited 2026 Aug 18]. Available from: https://ebooks.iospress.nl/doi/10.3233/SHTI220160 doi:10.3233/SHTI220160
23. Tweedy E, Mehta S, Mai MV, Ruiz V, Shi HL, Shaeffer GP, et al. Early prediction of unplanned critical care transfers in children using EHR-based ensemble machine learning. JAMIA open. 2026;9(3):ooag105.
24. Stein DF, Carter MJ, Booth J, Peters MJ, Ray S, Sebire NJ, et al. Predicting Cardiovascular deterioration in a paediatric intensive care unit (PicEWS): a machine learning modelling study of routinely collected health-care data. EClinicalMedicine [Internet]. 2025 [cited 2026 Aug 18];85. Available from: https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(25)00187-7/fulltext
25. van den Berg MA, O'Jay OAG, Benders MM, Bartels RR, Vijlbrief DD. Development and clinical impact assessment of a machine-learning model for early prediction of late-onset sepsis. Computers in Biology and Medicine. 2023;163:107156.

26. Shu CH, Zebda R, Espinosa C, Reiss J, Debuyserie A, Reber K, et al. Early prediction of mortality and morbidities in VLBW preterm neonates using machine learning. Pediatric research. 2025;97(6):2056–64.