Original Article

A nomogram to predict 28-day mortality in neonates with sepsis: a retrospective study based on the MIMIC-III database

Yongzhou Liang[#^], Liqing Zhao[#^], Jihong Huang[^], Yurong Wu[^]

Department of Pediatric Cardiology, Xinhua Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China

Contributions: (I) Conception and design: Y Wu, J Huang; (II) Administrative support: Y Wu, J Huang; (III) Provision of study materials or patients: J Huang, Y Liang; (IV) Collection and assembly of data: J Huang; (V) Data analysis and interpretation: Y Liang, L Zhao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work and should be considered as co-first authors.

Correspondence to: Yurong Wu, MD, PhD; Jihong Huang, PhD. Department of Pediatric Cardiology, Xinhua Hospital, School of Medicine, Shanghai Jiaotong University, 1665 Kongjiang Road, Yangpu District, Shanghai 200092, China. Email: wuyurong@xinhuamed.com.cn; huangjihong@xinhuamed.com.cn.

Background: Sepsis is the second-leading cause of death in neonates. We established a predictive nomogram to identify critically ill neonates early and reduce the time to treatment.

Methods: It is a retrospective case-control study based on the MIMIC-III database. The study population comprised 924 neonates diagnosed with sepsis.

Results: Neonates with sepsis included in the MIMIC-III database were enrolled, including 880 surviving neonates and 44 neonates who died. In the derivation dataset, stepwise regression and the Lasso algorithm were employed to select predictive variables, and the neonatal sequential organ failure assessment score (nSOFA) was calculated simultaneously. Bootstrap resampling was utilized to perform internal validation. The results indicated that the Lasso algorithm displayed superior discrimination, sensitivity, and specificity relative to stepwise regression and nSOFA scores. After 500 bootstrap resampling tests, the area under the receiver operating characteristic curve (AUC) of the Lasso algorithm was 0.912 (95% CI: 0.870–0.977). The nomogram based on the Lasso algorithm outperformed stepwise regression and nSOFA scores in terms of calibration and the clinical net benefit. This nomogram can assist in prognosticating neonatal severe sepsis and aid in guiding clinical practice while concurrently improving patient outcomes.

Conclusions: The established nomogram revealed that jaundice, corticosteroid use, weight, serum calcium, inotropes and base excess are all important predictors of 28-day mortality in neonates with sepsis. This nomogram can facilitate the early identification of neonates with severe sepsis. However, it still requires further modification and external validation to make it widely available.

Keywords: Sepsis; neonates; 28-day mortality; MIMIC database; nomogram

Submitted Mar 07, 2023. Accepted for publication Aug 02, 2023. Published online Sep 05, 2023.
doi: 10.21037/tp-23-150
View this article at: https://dx.doi.org/10.21037/tp-23-150

^ ORCID: Yongzhou Liang, 0000-0001-6782-6497; Liqing Zhao, 0000-0002-8268-4770; Jihong Huang, 0000-0001-7925-7244; Yurong Wu, 0000-0002-3311-2266.

Introduction

Neonatal death pertains to the death of live-born infants within the initial 28 days of life. Prematurity, neonatal sepsis and intrapartum-related complications are the most prevalent causes of neonatal death (1). Sepsis should be defined as life-threatening organ dysfunction resulting from a dysregulated host response to infection (2) and is perceived as the second-leading cause of death in neonates (3). A recent study on the global occurrence of neonatal sepsis indicated the incidence of 2,824 per 100,000 live births with a mortality rate up to 17.6% (4). Milton *et al.* reported an all-cause mortality of 5.65 per 1,000 neonate-days for those with clinically suspected and laboratory-confirmed neonatal sepsis (5). Despite the declining global case fatality rates of severe sepsis and septic shock in children over the past 30 years, the morbidity of neonatal sepsis is still underestimated due to the lack of a diagnostic gold standard (6). Sepsis carries a significant economic burden in low- and middle-income nations, amounting to over $7 billion in annual healthcare costs in the United States (7,8). Neonatal sepsis generally presents with diverse manifestations and non-specific laboratory results, which can rapidly progress to severe illness. Furthermore, the World Health Organization (WHO) recently accentuated that prevention can decrease sepsis morbidity and mortality (9). As such, the early detection and diagnosis of severe sepsis are imperative.

Currently, there are several prognostic tools available for identifying critical neonatal sepsis (10,11), including the Pediatric Logistic Organ Dysfunction score (PELOD), the Updated PELOD-2 score, Pediatric Risk of Mortality III scales (PRISM-III), Pediatric Multiple Organ Dysfunction score (P-MODS) and neonatal sequential organ failure assessment score (nSOFA). The PELOD score and PELOD-2 score included 12 and 10 factors, respectively, comprising five different organ dysfunctions, whereas the PRISM-III scales evaluated as many as 17 variables (12-14). These tools have limited operability and can be cumbersome, leading to time wastage and thus, are not applicable to newborns. Additionally, PELOD (or PELOD-2) shares commonalities with the P-MODS score and neglect to take into account the unique physiological traits of neonates and perinatal-related factors. Although nSOFA is simple and convenient, it only includes respiration, circulation and blood system score (15). Bestati *et al.* (16) has demonstrated that the individual correlation of neurological and hepatic dysfunction with neonatal mortality is absent in the evaluation of the nSOFA, resulting in a potential shortcoming. More recently, the machine learning (ML) algorithms have been utilized to forecast neonatal mortality in order to develop more effective tools (17), their restricted interpretability may hinder their usefulness in evaluating the prognosis of neonatal sepsis (18).

An urgent need exists for a straightforward and effective tool to predict neonatal sepsis mortality. Among the statistical visual tools available, nomogram has demonstrated concise and efficient performance when compared to traditional methods. With the evolution of statistical theory, nomograms have increasingly gained recognition and attention. Nomograms have indicated excellent performance in predicting late-onset sepsis in preterm infants with thyroid dysfunction (19) and have been widely used to determine the prognosis of critically ill patients in recent years (20,21). Currently, there is no relevant research that utilizes the nomogram as an instrument to foretell the mortality rate of neonatal sepsis to our understanding.

This study had dual objectives: firstly, to establish a nomogram by extracting neonatal demographic and comprehensive clinical data from the Medical Information Mart for Intensive Care III (MIMIC-III v1.4) database. Secondly, to compare the performance of the established nomogram with the nSOFA in predicting the mortality of neonates with sepsis within a 28-day period. Ultimately, the goal of our study is to establish a robust model to enable early identification of newborns at critically septic stages, which may help doctors to make informed treatment decisions based on a theoretical basis. We present this article in accordance with the TRIPOD reporting checklist (available at https://tp.amegroups.com/article/

Highlight box

Key findings

- Jaundice, weight, serum calcium, base excess, corticosteroids and inotropes were correlated with mortality in neonates with sepsis.

What is known and what is new?

- Sepsis is the second-leading cause of death in neonates, early identification and diagnosis of severe sepsis can improve neonatal outcomes.
- The nomogram will facilitate rapid identification of severe neonatal sepsis and will help clinicians make the right treatment decisions.

What is the implication, and what should change now?

- There is no accurate and concise method to identify neonates with severe sepsis, and it is necessary to develop and validate a tool for early identification of critically ill neonates.

view/10.21037/tp-23-150/rc).

Methods

Database and study population

This was a retrospective case-control study based on the MIMIC-III database. The MIMIC-III (v1.4) database is an open-access research database comprising information relating to the Intensive Care Unit (ICU) patients from the Beth Israel Deaconess Medical Center (Boston, MA, USA) between 2001 and 2012 (22). To extract data from the MIMIC-III database, we completed the National Institutes of Health's Protecting Human Research Participants (certification No. 44980177) web-based course.

Identity information and comprehensive clinical data of patients were integrated into the MIMIC-III database, which includes a total of 7,870 neonates. Since neonates were not directly involved, this study did not receive approval from the Institutional Review Committee board of Xinhua Hospital, Shanghai Jiaotong University, School of Medicine, Shanghai, China. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Massachusetts Institute of Technology (Cambridge, MA, USA) and Beth Israel Deaconess Medical Center (Boston, MA, USA), and individual consent for this retrospective analysis was waived.

Data extraction

Neonates who met the criteria for sepsis in the MIMIC-III database were included in this study. Sepsis was defined according to the Sepsis-3 criteria (2,23). Clinical data for the first clinical features of patients after admission were extracted from the MIMIC-III database using structured query language (SQL), including demographics, vital signs and laboratory tests. Comorbidities were identified using ICD-9 codes, including respiratory distress syndrome (RDS), jaundice, acute kidney injury (AKI) and others. AKI was diagnosed based on the Kidney Disease: Improving Global Outcomes (KDIGO) guidelines as an increase in serum creatinine (Scr) of 0.3 mg/dL or more (24). Data on medications, urine output and intravenous fluid volume within the first 24 h after ICU admission were extracted.

The primary outcome of this study was neonatal death, defined as death occurring within 28 d during hospitalization in the Neonatal Intensive Care Unit (NICU). In addition, the sequential organ failure assessment (SOFA) score of the neonates was extracted as a model. The nSOFA score uses categorical scores [total score ranging from 0 (best) to 15 (worst)] to objectively assess the condition of patients, with increasing scores indicating worse organ dysfunction (15,25), as follows: (I) receipt of mechanical ventilation and oxygen to maintain physiological peripheral saturation, with scores ranging from 0 to 8; (II) inotropic or vasoactive drug support, including the use of corticosteroids for presumed adrenal insufficiency or catecholamine-resistant shock, with scores ranging from 0 to 4; and (III) the presence and severity of thrombocytopenia based on the most recent platelet measure, with scores ranging from 0 to 3.

Missing data handling

Demographic information was complete, but missing data on laboratory tests and physiological items are common in the MIMIC-III database. However, if patients with incomplete data or incomplete variables are eliminated, a large bias would be produced. Therefore, in the study, patients or variables with more than 50% missing values were eliminated. Finally, the missing values in the database were replaced with the method "norm.predict" in Multivariate Imputation by Chained Equation (MICE), in which every variable is imputed conditional on all other variables (26).

Sample size and calculation

To precisely estimate the overall outcome risk or mean outcome value in our study, the required sample size was calculated as follows (27):

$$n = \left(1.96/\delta\right)^2 \cdot \hat{\varphi} \cdot \left(1-\hat{\varphi}\right) \quad [1]$$

The anticipated outcome proportion ($\hat{\varphi}$) in our study was 0.05, and the absolute margin of error (δ) was generally recommended to be less than 0.05. According to Eq. [1], the sample size required for the study was at least 73 participants.

To acquire a small prediction error in the estimated outcome probabilities across all individuals, the required sample size was calculated as follows (27):

$$\ln\left(MAPE\right) = -0.508 - 0.544\ln\left(n\right) + 0.259\ln\left(\varphi\right) + 0.504\left(P\right) \quad [2]$$

The mean absolute prediction error (MAPE) was no larger than 0.050. Moreover, 30 candidate predictor parameters (P) were extracted from the MIMIC-III database. According to Eq. [2], the sample size required for

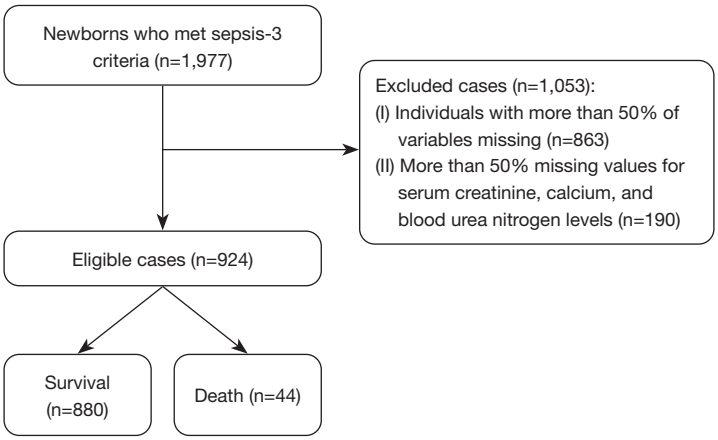

Figure 1 Flow chart for newborns screening process.

the study was at least 544 participants.

The study included neonates with sepsis who were admitted the NICU between 2001 and 2012, with an estimated sample size of 924.

Statistical analysis

Based on neonatal in-hospital death or survival as the dependent variable, we divided all neonates into two groups. Demographics, vital signs, laboratory tests, medications and comorbidities were defined as independent variables. According to the normality of distribution, continuous variables are presented as the mean ± SD for parametric variables or as the median and IQR for nonparametric variables. Categorical variables are expressed as numbers and percentages. For continuous variables, unpaired Student's test and Wilcoxon rank sum tests were used to assess significant differences between the two groups. For all categorical variables, the significant difference between the two groups was compared using the χ^2-test or Fisher's exact test.

Figure 1 shows the process of data acquisition, with 924 neonates included in our study. Based on the groupings described previously, we compared significant differences among the independent variables. Moreover, according to the dependent and independent variables described previously, a univariate logit binomial general linear model (GLM) was performed. The independent variables with $P<0.05$ in univariate analysis were incorporated into the GLM for stepwise regression analysis to further identify risk factors and establish a model (Further information regarding the model can be found in the Appendix 1 Supplement methods). The least absolute shrinkage and selection operator (LASSO) method was also used to screen variables, which can avoid overfitting by imposing a penalty on the magnitude of the model coefficients (28), and another model was established. To avoid multicollinearity among the variables, we employed both Pearson and nonparametric Spearman correlation matrices (29), with specific details presented in Figure S1. Additionally, the variance inflation factor (VIF) value of 5 or higher indicates the presence of multicollinearity (30). We examined potential nonlinear relationships between candidate continuous variables and death using restricted cubic splines and the Box-Tidwell test. Discrimination was assessed by the receiver operating characteristic (ROC) curve and area under the ROC curve (AUC), while calibration curves were plotted to assess model accuracy. Decision curve analysis (DCA) was also conducted to quantify the clinical net benefit at different threshold probabilities (31). Internal validation of the models and nSOFA score was performed using 500 bootstrap resamples and we compared the performance of the three models by calculating sensitivity, specificity, accuracy, integrated discrimination improvement (IDI), and net reclassification improvement (NRI). A nomogram was used to create a visual representation of the optimal model. This visualization tool includes graphic scoring which calculates the probability of a clinical event (32).

The SOFA (Sequential Organ Failure Assessment) score was originally developed to assess the severity of organ dysfunction or failure and was not designed to predict the risk of death (33). In recent years, the SOFA score has been extensively used to evaluate patient mortality (34,35), and the modified SOFA score (Table S1) has also been used

to assess the risk of death in newborns (15). The SOFA score was validated in this study and compared with our developed model in terms of discrimination, calibration and the clinical net benefit.

All statistical tests were two-sided, and P values ≤0.05 were considered to be statistically significant. We used R statistical software (V.4.1.1; https://www.R-project.org) and SPSS statistical software (SPSS Statistics 25.0) for statistical analyses.

Results

Baseline characteristics of the included neonates

A total of 1,977 neonates from the MIMIC-III database who met the Sepsis-3 criteria were included. A total of 1,053 neonates were excluded, including 863 patients with more than 50% missing variables and 190 patients with deficient creatinine, calcium and blood urea nitrogen levels. Table S2 provides comprehensive details on how the remaining missing values were handled. The selection procedure of the study population is shown in *Figure 1*. Ultimately, 924 patients were enrolled in our study, of whom 44 patients died and 880 patients were still alive after 28 days.

The comparisons of demographics and variables for the derivation dataset between patients who died and those who survived during hospitalization are shown in *Table 1*. The incidence of AKI was higher in the non-surviving group. Patients who died had lower serum bicarbonate, hematocrit, hemoglobin, RDW, calcium, bilirubin, base excess, weight, height, pO_2/FiO_2_mean and 24-hour urine levels. However, the levels of creatinine, chloride and pCO_2 in patients who died within 28 days were significantly increased. Additionally, more patients received systemic corticosteroids and inotropes in the non-surviving group. The nSOFA score of the non-surviving group was higher than that of the surviving group, which was consistent with previous reports. There was no significant difference in other variables between the surviving and non-surviving groups.

Selected features

Univariable and stepwise regression analysis was used to assess variables associated with death (*Table 2*). As shown in *Table 2*, stepwise regression analysis screened 5 candidate predictors with P values <0.05.

Figure 2 shows the results of 32 variables that were incorporated into the Lasso regression algorithm from the derivation group. Lasso regression was performed with a λ of 0.021 (one standard error of the minimum λ), and only 12 variables remained in the model. To make the model include fewer variables, lasso regression was performed with a λ of 0.028, and a total of 6 variables entered the model, which may be the most important candidate predictors for developing the model, including corticosteroid use, inotropes, jaundice, weight, serum calcium and base excess, as shown in *Table 3*.

Table 1 Comparison of the baseline characteristics of septic neonates with different outcomes

Characteristic	Overall (N=924)	Outcome		P
		Survival (N=880)	Death (N=44)	
Demography				
Gender, male	513 (55.5)	486 (55.2)	27 (61.4)	0.520
Age, hours	13.60 [8.10, 19.30]	13.62 [8.00, 19.31]	12.40 [9.93, 18.60]	0.968
Weight, kg	1.45 [1.04, 2.10]	1.46 [1.07, 2.11]	0.92 [0.69, 1.53]	<0.001
Height, cm	40.33±5.78	40.54±5.66	36.08±6.41	<0.001
Vital signs				
Heartrate, bmp	143.00±10.21	142.88±10.14	145.50±11.21	0.096
pO_2/FiO_2mean	200.00 [154.17, 277.52]	210.00 [159.09, 284.15]	131.35 [95.25, 197.60]	<0.001

Table 1 (*continued*)

Table 1 (*continued*)

Characteristic	Overall (N=924)	Outcome		P
		Survival (N=880)	Death (N=44)	
Laboratory tests				
Hematocrit, %	46.91±7.39	47.04±7.26	44.32±9.35	0.017
Hemoglobin, g/dL	15.60±2.45	15.64±2.42	14.78±2.92	0.023
Platelet, K/μL	249.33±84.03	250.36±83.97	228.82±83.58	0.097
RDW, %	16.90 [16.20, 17.80]	16.93 [16.22, 17.80]	16.15 [15.50, 17.12]	0.004
WBC, K/μL	9.30 [6.40, 13.40]	9.34 [6.46, 13.33]	10.20 [5.45, 13.57]	0.717
Neutrophils, K/μL	30.69±17.30	30.74±17.32	29.63±16.99	0.677
Lymphocytes, K/μL	56.80±18.54	56.73±18.50	58.14±19.56	0.624
Anion gap, mmol/L	16.33±3.63	16.28±3.50	17.36±5.57	0.053
Bicarbonate, mmol/L	20.94±3.16	20.99±3.10	19.91±4.09	0.027
Chloride, mmol/L	106.67±5.05	106.58±5.04	108.61±4.98	0.009
Calcium, mg/dL	9.68±1.39	9.76±1.33	8.08±1.56	<0.001
Serum potassium, mEq/L	4.89±1.05	4.90±1.05	4.86±0.99	0.859
Serum sodium, mEq/L	139.02±4.66	138.92±4.62	141.07±4.91	0.003
Bilirubin, mg/dL	4.93±1.95	5.03±1.92	3.02±1.35	<0.001
Creatinine, mg/dL	0.69±0.33	0.68±0.32	0.87±0.34	<0.001
BUN, mg/dL	17.00 [12.00, 26.00]	17.70 [12.00, 26.00]	18.00 [14.50, 24.25]	0.414
pCO_2, mmHg	47.90±11.90	47.60±11.35	54.00±19.15	<0.001
Base excess, mmol/L	−2.00 [−4.00, 0.00]	−2.10 [−3.98, −0.97]	−4.50 [−10.25, −2.00]	<0.001
Comorbidities				
Jaundice	798 (86.4)	779 (88.5)	19 (43.2)	<0.001
RDS	703 (76.1)	673 (76.5)	30 (68.2)	0.281
AKI (24 h)	57 (6.2)	47 (5.3)	10 (22.7)	<0.001
Others	381 (41.2)	364 (41.4)	17 (38.6)	0.840
Score				
nSOFA	4.00 [2.00, 6.00]	4.10 [2.10, 5.10]	7.00 [5.00, 9.25]	<0.001
Treatment				
Liquid intake, mL/d	131.78±54.67	132.39±53.77	119.60±70.09	0.130
Urine out, mL/d	86.93±47.55	89.16±47.13	42.18±31.38	<0.001
Corticosteroid	47 (5.1)	32 (3.6)	15 (34.1)	<0.001
Inotropes	123 (13.3)	95 (10.8)	28 (63.6)	<0.001

Data are presented as the mean ± SD, median [IQR], or numbers and percentages. RDW, red blood cell distribution width; WBC, white blood cell; BUN, blood urea nitrogen; RDS, respiratory distress syndrome; AKI, acute kidney injury; nSOFA, neonatal sequential organ failure assessment.

Table 2 Factors associated with death among septic neonates

Variables	Univariate analysis		Multivariate analysis	
	OR (95% CI)	P	OR (95% CI)	P
Demography				
Gender, male	1.30 (0.69, 2.40)	0.430		
Age, hours	1.00 (0.97, 1.01)	0.760		
Weight, kg	0.38 (0.22, 0.66)	0.001	0.47 (0.28, 0.81)	0.006
Height, cm	0.87 (0.82, 0.92)	0.000		
Vital signs				
Heartrate, bmp	1.00 (0.99, 1.10)	0.096		
pO_2/FiO_2mean	0.97 (0.96, 0.99)	0.000	0.98 (0.97, 0.99)	0.000
Laboratory tests				
Hematocrit, %	0.96 (0.93, 0.99)	0.017		
Hemoglobin, g/dL	0.88 (0.79, 0.98)	0.023		
Platelet, K/μL	1.00 (0.99, 1.01)	0.096		
RDW, %	0.85 (0.72, 1.00)	0.066		
WBC, K/μL	0.98 (0.94, 1.00)	0.520		
Neutrophils, K/μL	1.00 (0.98, 1.02)	0.680		
Lymphocytes, K/μL	1.00 (0.99, 1.01)	0.620		
Anion gap, mmol/L	1.10 (1.00, 1.20)	0.053		
Bicarbonate, mmol/L	0.90 (0.82, 0.99)	0.025	0.91 (0.81, 0.98)	0.027
Chloride, mmol/L	1.10 (1.00, 1.20)	0.009		
Calcium, mg/Dl	0.52 (0.43, 0.62)	0.000		
Serum potassium, mEq/L	0.97 (0.73, 1.30)	0.860		
Serum sodium, mEq/L	1.10 (1.00, 1.20)	0.003	1.09 (1.01, 1.16)	0.018
Bilirubin, mg/dL	0.43 (0.34, 0.55)	0.000		
Creatinine, mg/dL	3.70 (1.80, 7.60)	0.000	4.38 (1.91, 10.05)	0.001
BUN, mg/dL	1.00 (0.98, 1.01)	0.600		
pCO_2, mmHg	1.00 (0.91, 1.10)	0.001		
Base excess, mmol/L	0.86 (0.82, 0.91)	0.000		
Comorbidities				
Jaundice	0.10 (0.05, 0.19)	0.000		
RDS	0.66 (0.34, 1.30)	0.210		
AKI (24 h)	5.20 (2.40, 11.00)	0.000		
Others	0.89 (0.48, 1.70)	0.720		
Treatment				
Liquid intake, mL/d	0.99 (0.99, 1.00)	0.130		
Urine out, mL/d	0.96 (0.95, 0.97)	0.000		
Corticosteroid	14.00 (6.70, 28.00)	0.000		
Inotropes	14.00 (7.50, 28.00)	0.000		

RDW, red blood cell distribution width; WBC, white blood cell; BUN, blood urea nitrogen; RDS, respiratory distress syndrome; AKI, acute kidney injury; nSOFA, neonatal sequential organ failure assessment.

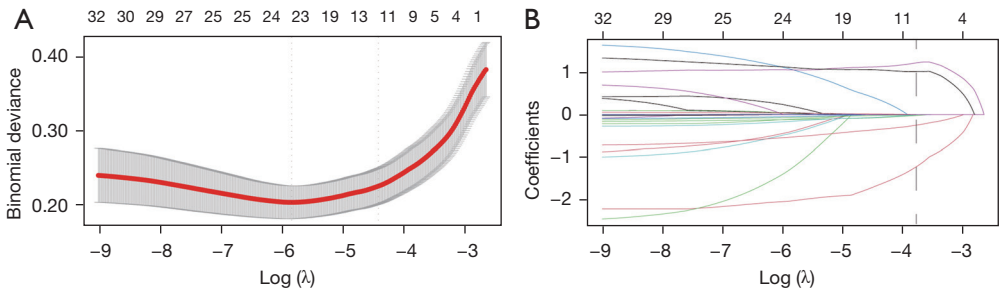

Figure 2 Variables selection by the LASSO regression. (A) The tuning parameter (λ) in the LASSO model was selected by 10-fold cross-validation via minimum criteria. The partial likelihood deviance (binomial deviance) curve was plotted versus log(λ). The results showed that 23 variables were retained when the error was the smallest, which corresponded to the dotted line on the left. To avoid overfitting and simplicity of the model, 12 variables were retained, corresponding to the dotted line on the right, which is the one SE of the minimum criteria (the 1–SE criteria). To make the model simpler, lasso regression was performed with λ of 0.028, and a total of 6 variables enter the model. (B) LASSO coefficient profiles of the 32 variables. A coefficient profile plot was produced against the log(λ) sequence. A vertical line was drawn at the selected optimizing value(λ), corresponding to 6 nonzero coefficients. LASSO, the least absolute shrinkage and selection operator; SE, standard error.

Table 3 LASSO selected predictors

Variables	LASSO selected predictors	
	OR (95% CI)	P
Jaundice	0.05 (0.02, 0.14)	0.001
Corticosteroid	3.81 (1.35, 10.72)	0.011
Weight, kg	0.22 (0.11, 0.43)	0.001
Calcium, mg/dL	0.53 (0.40, 0.69)	0.001
Inotropes	3.58 (1.50, 8.55)	0.004
Base excess, mmol/L	0.94 (0.88, 0.99)	0.048

LASSO, the least absolute shrinkage and selection operator.

Model development

For the derivation dataset, candidate variables obtained by lasso regression that had a significant effect on inhospital death are shown in *Table 3*. In addition, candidate variables with statistically significant differences in stepwise regression analysis were included in the binary logistic regression, as shown in *Table 2*. Ultimately, the following 6 candidate factors, identified as statistically significant variables using the multivariate likelihood ratio (LR), were included in the Lasso regression algorithm: jaundice (OR 0.05, P=0.001), corticosteroid use (OR 3.81, P=0.011), weight (OR 0.22, P=0.001), serum calcium (OR 0.53, P=0.001), inotropes (OR 3.58, P=0.004) and base excess (OR 0.94, P=0.048). Correspondingly, the binary logistic regression model incorporated a total of 5 variables from the derivation dataset, which included weight (OR 0.47, P=0.006), pO_2/FiO_2_mean (OR 0.98, P=0.000), bicarbonate (OR 0.91, P=0.027), serum sodium (OR 1.09, P=0.018) and creatinine (OR 4.38, P=0.001). Ultimately, both the lasso and binary logistic regression models exhibited no nonlinearity relationships in the continuous variables, as depicted in *Figure 3*. A linear correlation was observed between nSOFA score and mortality. The candidate variables did not exhibit multicollinear relationships (Figure S1), and the VIFs are presented in *Table 4*. We developed two regression models, the lasso regression model and binary logistic regression model, using the candidate variables screened earlier.

Validation and comparison of the models

The discrimination of the two models and the nSOFA score in the derivation dataset was assessed using the ROC curve. In the derivation dataset, the AUC values of three prediction models, namely the stepwise regression model, Lasso algorithm model, and nSOFA score, were determined as 0.784 (95% CI: 0.703–0.866), 0.924 (95% CI: 0.869–0.978), and 0.807 (95% CI: 0.728–0.866), respectively (Table S3). Following 500-times bootstrap resampling, the values changed to 0.763 (95% CI: 0.695–0.874), 0.912 (95% CI: 0.870–0.977), and 0.807 (95% CI: 0.732–0.883) (*Figure 4*). The discrimination of the LASSO algorithm model was better than that of the stepwise regression

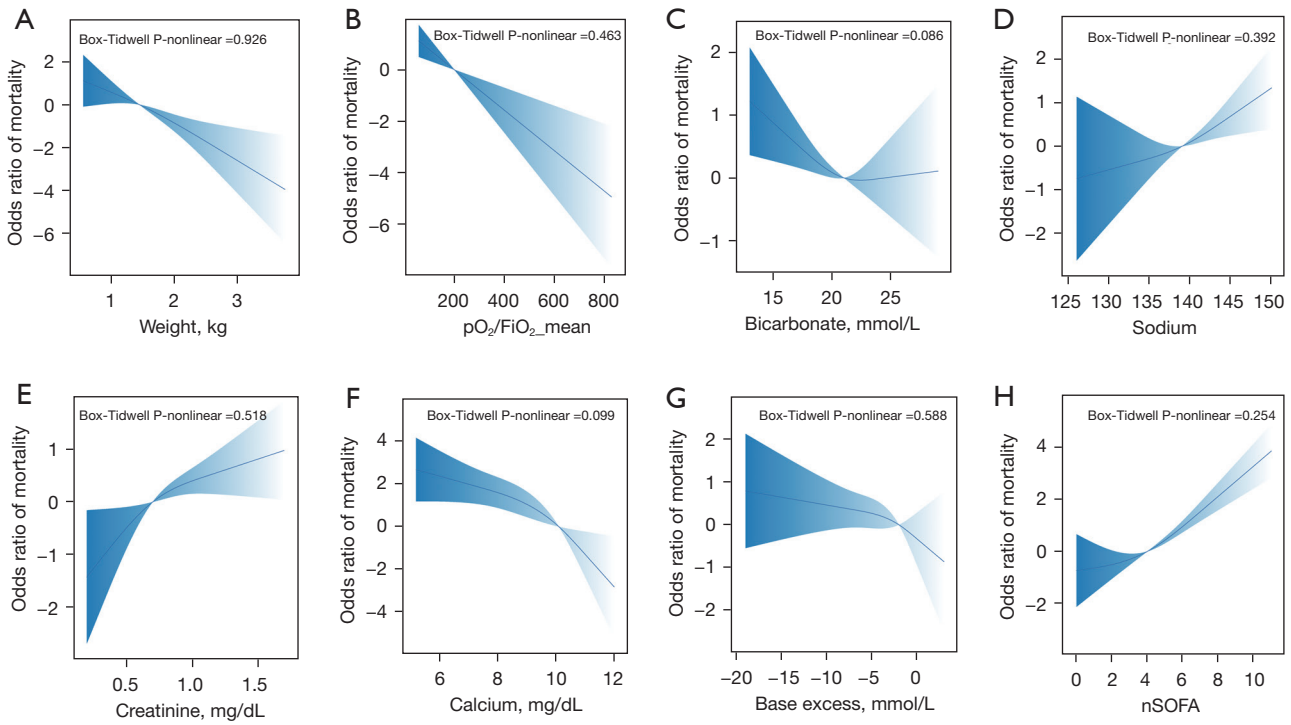

Figure 3 Association of predicted variables with mortality. (A) Weight. (B) pO_2/FiO_2_mean. (C) Bicarbonate. (D) Sodium. (E) Creatinine. (F) Calcium. (G) Base excess. (H) nSOFA. Odds ratios are indicated by solid lines and 95% CIs by shaded areas; Box-Tidwell P-nonlinear <0.05 indicates a nonlinear relationship; nSOFA, neonatal sequential organ failure assessment score.

Table 4 VIFs for stepwise regression and LASSO regression

Variables	VIF
Multivariate analysis	
Bicarbonate, mmol/L	1.011
Serum sodium, mEq/L	1.066
Weight, kg	1.058
pO_2/FiO_2mean	1.003
Creatinine, mg/dL	1.009
LASSO selected predictors	
Jaundice	1.332
Corticosteroid	1.138
Weight, kg	1.429
Calcium, mmol/L	1.213
Inotropes	1.208
Base excess, mmol/L	1.195

VIFs, variance inflation factors; LASSO, the least absolute shrinkage and selection operator.

model and nSOFA score (P<0.05) (*Figure 5*). There was no statistically significant difference in the discrimination performance between the stepwise regression model and the nSOFA score (P=0.418) (*Figure 5C*). In addition, the sensitivity, specificity and accuracy of each model are presented in Table S3. The stepwise regression model and nSOFA score displayed high sensitivity and accuracy, but their specificity was lower. In contrast, the Lasso algorithm model demonstrated better sensitivity, specificity, and accuracy. Moreover, the calibration curve plot of the Lasso algorithm model and nSOFA score model roughly showed an ideal diagonal in the derivation dataset (*Figure 4D*,*4F*). Instead, the stepwise regression model exhibited a relatively low consistency among the models (*Figure 4B*). Decision curve analysis (DCA) was applied to evaluate the clinical net benefit across the entire range of threshold probabilities for each of the three models (*Figure 5D*). Results indicate that the Lasso algorithm model had the highest clinical and public health value.

The performance of several models predicting mortality

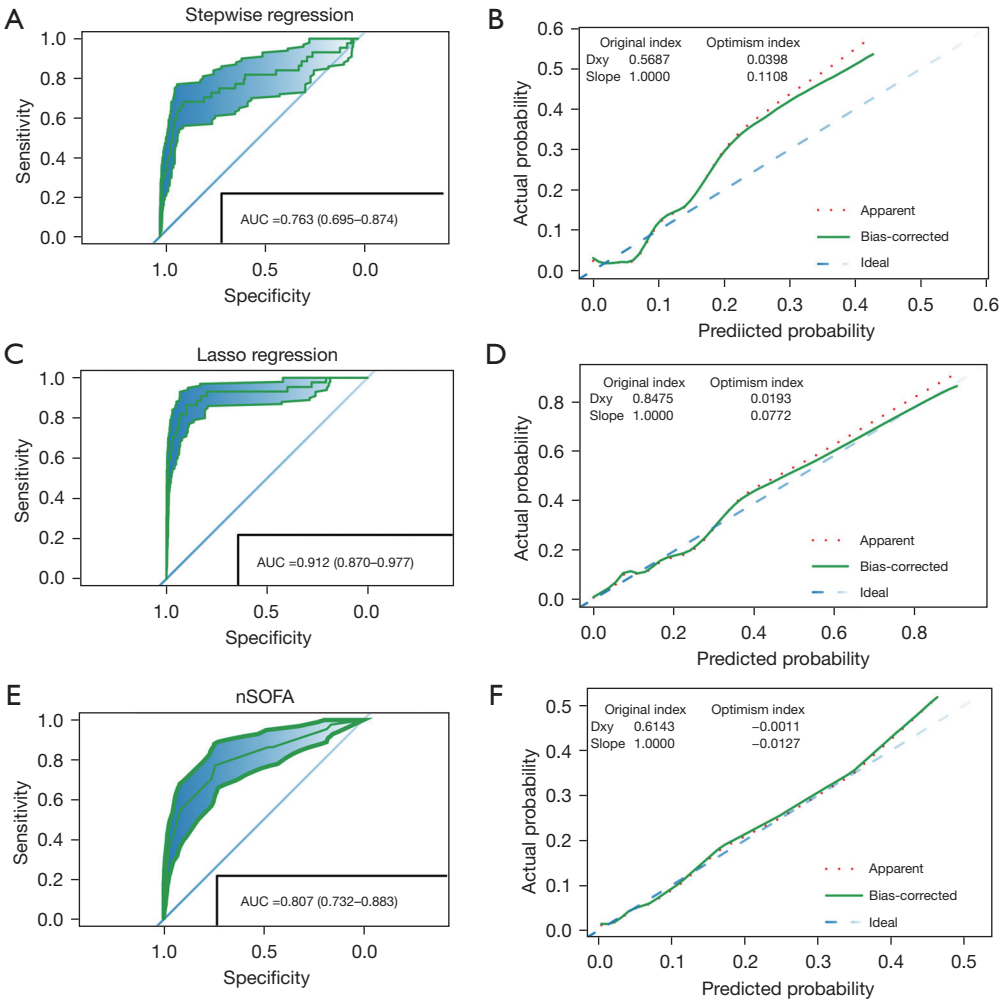

Figure 4 ROC curves and calibration curve for predicting mortality of sepsis neonates. (A) ROC curve for Stepwise regression. (B) Calibration curve for Stepwise regression. (C) ROC curve for LASSO. (D) Calibration curve for LASSO. (E) ROC curves for nSOFA. (F) Calibration curve for nSOFA. Slope, calibration slope; Optimism index derives from internal validation; ROC, receiver operating characteristic; AUC, area under the ROC curve; LASSO, the least absolute shrinkage and selection operator; nSOFA, neonatal sequential organ failure assessment score.

in neonatal populations with sepsis was compared (Table S4). The Lasso algorithm significantly enhances prediction performance compared to that of the stepwise regression and nSOFA score. The net reclassification improvement (NRI) of Lasso algorithm with respect to stepwise regression and nSOFA score are 52.95% and 53.64%, respectively. Moreover, Lasso algorithm improved model performance validated by integrated discrimination improvement (IDI) by 30.93% and 26.98% compared to stepwise regression and nSOFA score, respectively. Overall, the Lasso algorithm model outperforms both the stepwise regression model and nSOFA score in terms of performance.

Development of the best-performing prediction nomogram

The 6 factors had nonzero coefficients in the Lasso regression model based on the derivation groups (*Table 3*), which were integrated into the nomogram (R2 =0.526, C-index =0.912). A candidate predictor corresponds to a unique score and then sums these scores to obtain a total score, which corresponds to the total score axis and finally to the inhospital mortality axis (*Figure 6*). Physicians can

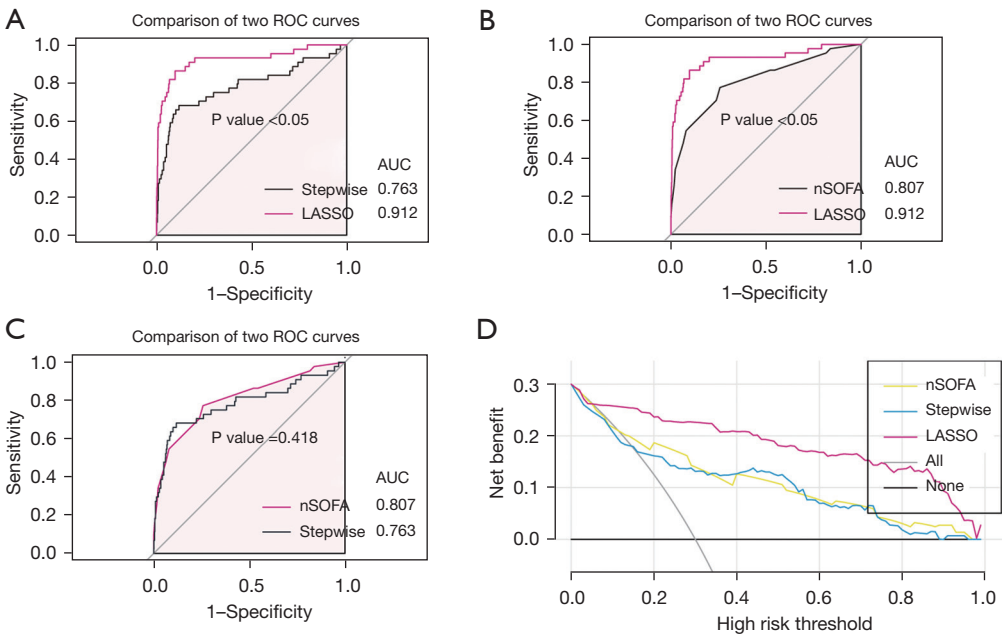

Figure 5 Comparison of the discrimination and decision curves of the stepwise regression, LASSO and nSOFA. (A) the stepwise regression and LASSO. (B) the nSOFA and LASSO. (C) the stepwise regression and nSOFA. (D) Comparison of the net benefit among the stepwise regression, LASSO and nSOFA. LASSO, the least absolute shrinkage and selection operator; nSOFA, neonatal sequential organ failure assessment; ROC, receiver operating characteristic; AUC, area under the ROC curve.

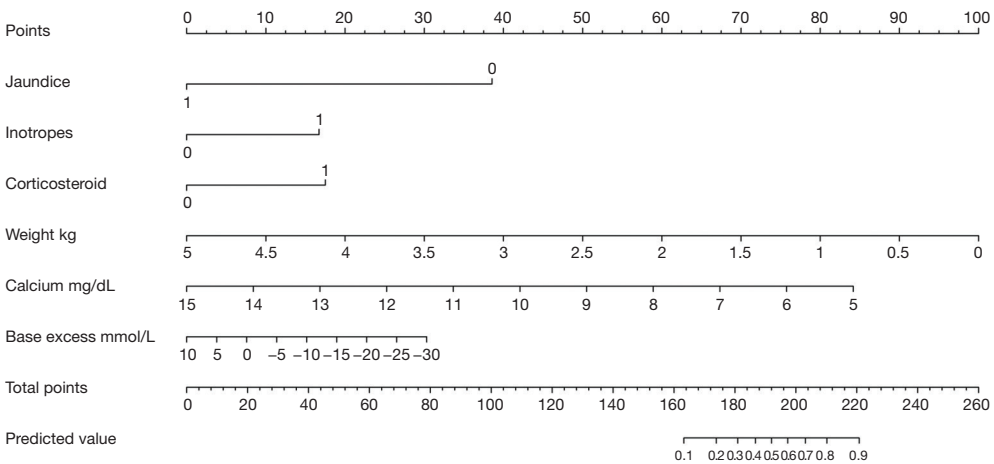

Figure 6 The nomogram for predicting 28-day in-hospital mortality of sepsis neonates.

make the corresponding treatment decisions according to the probability of death that is obtained.

Discussion

Sepsis is a major cause of neonatal death; the mortality of sepsis can be reduced by identifying critically ill patients early (3). Because the clinical signs at the early stages of neonatal sepsis are non-specific, there is currently no uniform standard for early identification of critically ill patients, resulting in delayed management of critically ill patients, which may negatively affect their clinical outcomes. Therefore, based on data derived from the MIMIC-III database, we selected independent risk factors for mortality

of neonates with sepsis by using stepwise regression and LASSO regression analysis. Meanwhile, two mortality prediction models with independent predictors for neonates with sepsis were developed. The optimal model was selected by comparing the two prediction models with the nSOFA score model. Finally, a neonatal sepsis mortality prediction nomogram was developed based on the optimal model. The nomogram may facilitate the early identification of neonates with severe sepsis and may reduce mortality of critically ill patients.

In 2015, neonatal deaths accounted for approximately 45.1% of children who did not live to 5 years of age, with one of the leading causes being sepsis or meningitis (1). A study showed mortality ranging from 11% in the USA to 19% in India (3). Another recent study showed that the all-cause mortality of neonatal sepsis was 0.83 (0.37–2.00) per 1,000 neonate-days in low-income and middle-income countries (5). The mortality in our study was 5%, which was lower than that previously described. These differences in mortality may be attributed to better medical conditions in the United States and relatively advanced neonatal care.

Despite advances in neonatal intensive care technology, nonspecific clinical symptoms often lead to rapidly progress to severe illness. Therefore, we developed a mortality prediction model to provide a simpler tool for physicians to identify neonates with severe sepsis early that may help reduce the mortality of infection-related illnesses. Our results indicated that weight, pO_2/FiO_2, serum calcium, base excess and bicarbonate had protective effects on neonatal in-hospital death, while systemic corticosteroids, inotropes, and high serum creatinine levels were risk factors within the first 24 h after NICU admission. Previous studies have demonstrated that weight, pO_2/FiO_2, serum calcium, base excess and bicarbonate are protective factors associated with neonatal death (15,36-38). Firstly, the pO_2/FiO_2 is an available noninvasive marker for the identification of children with acute lung injury or ARDS (39). Secondly, serum calcium is crucial for various physiological processes, including signal transduction in cells, nerve transmission, coagulation cascade and muscle contraction (40). Neonatal sepsis frequently causes hypocalcemia due to the presence of inflammatory markers, including procalcitonin (PCT), tumor-necrosis factor α (TNF-α), interleukin-1β (IL-1β) (41). Further research demonstrates that Neonatal sepsis accompanied by hypocalcemia increases the risk for developing cardiovascular and renal dysfunction, as well as disseminated intravascular coagulation and epilepsy, ultimately resulting in higher mortality rates (42). Additionally, animal study supports the involvement of calcium ions in lethal coagulation during sepsis (43). Thirdly, hemodynamic instability during neonatal sepsis results in tissue hypoperfusion, causing metabolic acidosis and consequently, reduced levels of serum bicarbonate and BE. Bicarbonate effectively improves tissue perfusion, myocardial contractility, and cellular dysfunction in severe acidosis (44,45). Reduced levels of bicarbonate are known to correlate with decreased levels of BE. Our study did not utilize lactate to assess tissue hypoperfusion, as excessive lactate deficiency can lead to unreliable results. A study reveals that increase of negative value of BE values can better reflect hypoperfusion-induced metabolic acidosis in critically ill patients compared to lactate (46). These findings also offset certain limitations in our study. Moreover, low serum levels of both bicarbonate and BE are indicative of a higher mortality in critically ill patients (46-49). These studies support our conclusions while further enhancing our understanding of the association between these indicators and neonatal mortality.

The literature has presented an elevation of plasma bilirubin levels indicates strong liver conversion function, while lower concentration of cholestasis-related markers implies potential protective effect of bilirubin (50). Previous study has indicated that bilirubin accumulation may be beneficial for the survival of newborns and septic neonates who survived had higher plasma bilirubin levels (51). Stocker *et al.* believed that Bilirubin is a strong antioxidant whose ability to suppress the oxidation *in vitro* is more prominent than that of another powerful antioxidant-α-tocophero (52). Also, bilirubin can protect low-density lipoprotein from damage caused by oxidation (53). Meanwhile, for newborns suffering from sepsis, inhibiting GBS growth is possible by affecting substrate utilization when the plasma bilirubin concentration is below the diagnostic criteria for neonatal hyperbilirubinemia (54). These results demonstrated a protective effect of jaundice on neonatal sepsis. However, the mechanism by which bilirubin provides such protection against neonatal sepsis remains unclear and requires additional research.

Our study found that systematic corticosteroids and inotropes were associated with adverse outcomes in neonates, supported by the nSOFA score (15). The efficacy of corticosteroids can improve mortality in patients with sepsis remains controversial, however, the administration of corticosteroids was associated with more superinfection, hyperglycemia and hypernatremia (55). Additionally, corticosteroids are primarily prescribed to children with sepsis who suffer from fluid refractory and catecholamine

resistant shock (56). Simultaneously, refractory septic shock in children was associated with high vaso-inotrope doses, indicating that neonates had more critical outcomes (57). The inotropes represented circulatory instability and it was a major cause of arrhythmia and leaded to some life-threatening complications such as central nervous bleeding and intestinal ischemia (58). Furthermore, high serum creatinine levels indicated worsening acute kidney injury associated with poor outcomes (59,60), which confirmed our conclusion.

Among the available sickness prediction scoring systems for assessing the severity of neonates, the Clinical Risk Index for Babies (CRIB) and CRIB-II scores primarily evaluate several parameters during the perinatal period and are applicable to newborns with gestational age <33 weeks (61). In addition, The Transport Risk Index of Physiologic Stability (TRIPS) and TRIPS-II scores evaluate temperature, respiratory status, blood pressure, response to noxious stimuli in infants (62,63). TRIPS and TRIPS-II scores are applied to rapidly assess the infants requiring transport to a NICU, and cannot accurately assess the severity of neonatal sepsis. The Score for Neonatal Acute Physiology (SNAP) consists of 27 items of vital signs, blood gas indicators, peripheral blood cells, and biochemistry indicators (64). The SNAP Perinatal Extension (SNAPPE), derived from SNAP, includes three supplementary indicators, that is, birth weight, small for gestational age (SGA), and Apgar score at 5 min. However, their applicability is limited by the complexity of the calculations and the multiple variables involved. The simplified SNAP-II score includes mean blood pressure, lowest body temperature, $P(O_2)/FiO_2$, pH, urine output, and multiple seizures, but its performance in evaluating the severity of neonatal diseases is not ideal (65). The SNAPPE-II score adds birth weight, SGA, and Apgar score at 5 min to the SNAP-II score. Previous studies compared the CRIB-II and SNAPPE-II for predicting mortality, and neither of them fully estimated the risk of death (66). Obviously, an accurate scoring system with fewer variables is more convenient for clinical application than relying on multiple variables. In our study, the nomogram incorporates only 6 easily ascertainable variables, while considering the impact of organ function and treatment measures on the progression of sepsis. These are factors that CRIB-II, SNAP-II, and SNAPPE-II scores, commonly used to assess the severity of neonates, have not considered.

The optimal SOFA cutoff to discriminate mortality was a score higher than 8 points, and the performance of the maximum SOFA score in discriminating in-hospital mortality was similar to the performance of the PRISM and PELOD-2 and better than the P-MODS score (10). The PRISM, PELOD-2, and P-MODS have been validated in the prognosis of sepsis in children (67) but have not been applied to neonates. Moreover, a SOFA score of 2 or higher has limited sensitivity in predicting in-hospital mortality risk (11). However, our findings indicate that nSOFA has relatively high sensitivity but low specificity. We noticed several relevant parallels of our study with studies on SOFA performance (*Figure 3*). As the SOFA score increased, the risk of death increased (10,11,15), and the SOFA performance (AUC, 0.81; 95% CI: 0.73–0.88) was similar to the conclusion of a previous study (AUC, 0.81; 95% CI: 0.76–0.85) (15), which proved that the performance of the nomogram may be better than that of the SOFA score.

As prognostic devices, nomograms have been widely used to predict the probability of clinical events. We developed a nomogram for predicting the risk of in-hospital mortality in neonates, which can enable physicians to quickly identify severe sepsis in neonates and make rapid clinical decisions. Moreover, the nomogram showed good predictive accuracy in predicting mortality in neonates due to severe sepsis, with an AUC of 0.912. However, there are some inevitable limitations in our study that need to be addressed. First, the study was a single-center retrospective study based on the MIMIC-III database, which made it impossible for external validation. Second, the model's statistical power was affected by a small sample size, as there were fewer neonates in the non-surviving group. Although we utilized a recent technique for calculating sample size, the conventional method that meets Peduzzi's criterion of an event per variable (EPV) of >10 shows that we cannot achieve adequate statistical power. In addition, the etiology of individual patient-level information were not available. This may limit our more accurate conclusion. Moreover, due to the late availability of etiological information, its role in accurate medication within 24 h of admission was relatively limited. Nonetheless, our findings provide novel insights into the prediction of mortality in neonates with sepsis, which has potentially important public health implications for improving survival of neonates with sepsis. Our study's limitations should be acknowledged, that need to be addressed in future studies.

Conclusions

Jaundice, weight, serum calcium and base excess were

associated with a reduced risk of mortality in neonates with sepsis, while corticosteroids and inotropes were related to an increased risk of mortality in neonates with sepsis. We developed a mortality prediction nomogram with a small number of routinely collected variables for neonates with sepsis, which can be applied to identify critically ill neonates early. The nomogram can be easily used in the NICU.

Acknowledgments

We thank all participants in the Xinhua Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China. The authors gratefully acknowledge the editors and reviewers of Transnational Pediatrics for their invaluable feedback, particularly their insightful comments on statistical issues that improved the manuscript. Additionally, we appreciate their support in editing our manuscript.
Funding: This work was supported in part by grants from the National Natural Science Foundation of China (NSFC Project No. 82171948), and the Shanghai Municipal Health Commission Scientific Research Project (Project No. 20214Y0126). The sponsors of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tp.amegroups.com/article/view/10.21037/tp-23-150/rc

Peer Review File: Available at https://tp.amegroups.com/article/view/10.21037/tp-23-150/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tp.amegroups.com/article/view/10.21037/tp-23-150/coif). All authors report that this work was supported in part by grants from the National Natural Science Foundation of China (NSFC Project No. 82171948), and the Shanghai Municipal Health Commission Scientific Research Project (Project No. 20214Y0126). The sponsors of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Massachusetts Institute of Technology (Cambridge, MA) and Beth Israel Deaconess Medical Center (Boston, MA), and individual consent for this retrospective analysis was waived.

References

1. Liu L, Oza S, Hogan D, 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 2016;388:3027-35.
2. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016;315:801-10.
3. Fleischmann-Struzek C, Goldfarb DM, Schlattmann P, et al. The global burden of paediatric and neonatal sepsis: a systematic review. Lancet Respir Med 2018;6:223-30.
4. Fleischmann C, Reichert F, Cassini A, et al. Global incidence and mortality of neonatal sepsis: a systematic review and meta-analysis. Arch Dis Child 2021. [Epub ahead of print]. doi: 10.1136/archdischild-2020-320217.
5. Milton R, Gillespie D, Dyer C, 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;10:e661-72.
6. Tan B, Wong JJ, Sultana R, et al. Global Case-Fatality Rates in Pediatric Severe Sepsis and Septic Shock: A Systematic Review and Meta-analysis. JAMA Pediatr 2019;173:352-62.
7. Ranjeva SL, Warf BC, Schiff SJ. Economic burden of neonatal sepsis in sub-Saharan Africa. BMJ Glob Health 2018;3:e000347.
8. Carlton EF, Barbaro RP, Iwashyna TJ, et al. Cost of

Pediatric Severe Sepsis Hospitalizations. JAMA Pediatr 2019;173:986-7.
9. Board WHOE. EB140/12: Improving the prevention, diagnosis and clinical management of sepsis Published 2017. Available online: http://apps.who.int/gb/ebwha/pdf_files/EB140/B140_12-en.pdf.
10. Matics TJ, Sanchez-Pinto LN. Adaptation and Validation of a Pediatric Sequential Organ Failure Assessment Score and Evaluation of the Sepsis-3 Definitions in Critically Ill Children. JAMA Pediatr 2017;171:e172352.
11. Balamuth F, Scott HF, Weiss SL, et al. Validation of the Pediatric Sequential Organ Failure Assessment Score and Evaluation of Third International Consensus Definitions for Sepsis and Septic Shock Definitions in the Pediatric Emergency Department. JAMA Pediatr 2022;176:672-8.
12. Leteurtre S, Martinot A, Duhamel A, et al. Validation of the paediatric logistic organ dysfunction (PELOD) score: prospective, observational, multicentre study. Lancet 2003;362:192-7.
13. Leteurtre S, Duhamel A, Salleron J, et al. PELOD-2: an update of the PEdiatric logistic organ dysfunction score. Crit Care Med 2013;41:1761-73.
14. Graciano AL, Balko JA, Rahn DS, et al. The Pediatric Multiple Organ Dysfunction Score (P-MODS): development and validation of an objective scale to measure the severity of multiple organ dysfunction in critically ill children. Crit Care Med 2005;33:1484-91.
15. Fleiss N, Coggins SA, Lewis AN, et al. Evaluation of the Neonatal Sequential Organ Failure Assessment and Mortality Risk in Preterm Infants With Late-Onset Infection. JAMA Netw Open 2021;4:e2036518.
16. Bestati N, Leteurtre S, Duhamel A, et al. Differences in organ dysfunctions between neonates and older children: a prospective, observational, multicenter study. Crit Care 2010;14:R202.
17. Hsu JF, Chang YF, Cheng HJ, et al. Machine Learning Approaches to Predict In-Hospital Mortality among Neonates with Clinically Suspected Sepsis in the Neonatal Intensive Care Unit. J Pers Med 2021;11:695.
18. Yoon CH, Torrance R, Scheinerman N. Machine learning in medicine: should the pursuit of enhanced interpretability be abandoned? J Med Ethics 2022;48:581-5.
19. Huang Y, Yu X, Li W, et al. Development and validation of a nomogram for predicting late-onset sepsis in preterm infants on the basis of thyroid function and other risk factors: Mixed retrospective and prospective cohort study. J Adv Res 2020;24:43-51.
20. Lu Z, Zhang J, Hong J, et al. Development of a Nomogram to Predict 28-Day Mortality of Patients With Sepsis-Induced Coagulopathy: An Analysis of the MIMIC-III Database. Front Med (Lausanne) 2021;8:661710.
21. Hou N, Li M, He L, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. J Transl Med 2020;18:462.
22. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016;3:160035.
23. Weiss SL, Peters MJ, Alhazzani W, et al. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. Intensive Care Med 2020;46:10-67.
24. Kellum JA, Lameire N; . Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). Crit Care 2013;17:204.
25. Lambden S, Laterre PF, Levy MM, et al. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. Crit Care 2019;23:374.
26. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw 2011;45:1-67.
27. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368:m441.
28. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Stat Med 2007;26:5512-28.
29. Fabio A, Li W, Strotmeyer S, et al. Racial segregation and county level intentional injury in Pennsylvania: analysis of hospital discharge data for 1997-1999. J Epidemiol Community Health 2004;58:346-51.
30. Shen Y, Huang X, Zhang W. Platelet-to-lymphocyte ratio as a prognostic predictor of mortality for sepsis: interaction effect with disease severity-a retrospective study. BMJ Open 2019;9:e022896.
31. Kerr KF, Brown MD, Zhu K, et al. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. J Clin Oncol 2016;34:2534-40.
32. Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008;26:1364-70.
33. Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care

Medicine. Crit Care Med 1998;26:1793-800.
34. Raith EP, Udy AA, Bailey M, et al. Prognostic Accuracy of the SOFA Score, SIRS Criteria, and qSOFA Score for In-Hospital Mortality Among Adults With Suspected Infection Admitted to the Intensive Care Unit. JAMA 2017;317:290-300.
35. Elias A, Agbarieh R, Saliba W, et al. SOFA score and short-term mortality in acute decompensated heart failure. Sci Rep 2020;10:20802.
36. Park HW, Park SY, Kim EA. Prediction of In-Hospital Mortality After 24 Hours in Very Low Birth Weight Infants. Pediatrics 2021;147:e2020004812.
37. Zhang K, Zhang S, Cui W, et al. Development and Validation of a Sepsis Mortality Risk Score for Sepsis-3 Patients in Intensive Care Unit. Front Med (Lausanne) 2020;7:609769.
38. Knutzen L, Svirko E, Impey L. The significance of base deficit in acidemic term neonates. Am J Obstet Gynecol 2015;213:373.e1-7.
39. Khemani RG, Patel NR, Bart RD 3rd, et al. Comparison of the pulse oximetric saturation/fraction of inspired oxygen ratio and the PaO2/fraction of inspired oxygen ratio in children. Chest 2009;135:662-8.
40. Jia X, Chen X, Gao C, et al. Functional cooperation between IK(Ca) and TRPC1 channels regulates serum-induced vascular smooth muscle cell proliferation via mediating Ca(2+) influx and ERK1/2 activation. Cell Prolif 2023;56:e13385.
41. Kutílek Š, Vracovská M, Pečenková K, et al. Calcemia and Inflammatory Markers in Early-Onset Neonatal Infection. Acta Medica (Hradec Kralove) 2019;62:58-61.
42. Liu Y, Chai Y, Rong Z, et al. Prognostic Value of Ionized Calcium Levels in Neonatal Sepsis. Ann Nutr Metab 2020;76:193-200.
43. Zhang H, Zeng L, Xie M, et al. TMEM173 Drives Lethal Coagulation in Sepsis. Cell Host Microbe 2020;27:556-570.e6.
44. Kraut JA, Madias NE. Lactic acidosis. N Engl J Med 2014;371:2309-19.
45. Berend K, de Vries AP, Gans RO. Physiological approach to assessment of acid-base disturbances. N Engl J Med 2014;371:1434-45.
46. Yuan J, Liu X, Liu Y, et al. Association between base excess and 28-day mortality in sepsis patients: A secondary analysis based on the MIMIC- IV database. Heliyon 2023;9:e15990.
47. Couto-Alves A, Wright VJ, Perumal K, et al. A new scoring system derived from base excess and platelet count at presentation predicts mortality in paediatric meningococcal sepsis. Crit Care 2013;17:R68.
48. Tan L, Xu Q, Li C, et al. Association Between the Admission Serum Bicarbonate and Short-Term and Long-Term Mortality in Acute Aortic Dissection Patients Admitted to the Intensive Care Unit. Int J Gen Med 2021;14:4183-95.
49. Allen CJ, Wagenaar AE, Horkan DB, et al. Predictors of mortality in pediatric trauma: experiences of a level 1 trauma center and an assessment of the International Classification Injury Severity Score (ICISS). Pediatr Surg Int 2016;32:657-63.
50. Jenniskens M, Langouche L, Vanwijngaerden YM, et al. Cholestatic liver (dys)function during sepsis and other critical illnesses. Intensive Care Med 2016;42:16-27.
51. Khalil S, Shah D, Faridi MM, et al. Prevalence and outcome of hepatobiliary dysfunction in neonatal septicaemia. J Pediatr Gastroenterol Nutr 2012;54:218-22.
52. Stocker R, Yamamoto Y, McDonagh AF, et al. Bilirubin is an antioxidant of possible physiological importance. Science 1987;235:1043-6.
53. Tomaro ML, Batlle AM. Bilirubin: its role in cytoprotection against oxidative stress. Int J Biochem Cell Biol 2002;34:216-20.
54. Hansen R, Gibson S, De Paiva Alves E, et al. Adaptive response of neonatal sepsis-derived Group B Streptococcus to bilirubin. Sci Rep 2018;8:6470.
55. Chan ED, Chan MM, Chan MM, et al. Use of glucocorticoids in the critical care setting: Science and clinical evidence. Pharmacol Ther 2020;206:107428.
56. Rhodes A, Evans LE, Alhazzani W, et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Intensive Care Med 2017;43:304-77.
57. Morin L, Ray S, Wilson C, et al. Refractory septic shock in children: a European Society of Paediatric and Neonatal Intensive Care definition. Intensive Care Med 2016;42:1948-57.
58. Annane D, Ouanes-Besbes L, de Backer D, et al. A global perspective on vasoactive agents in shock. Intensive Care Med 2018;44:833-46.
59. Choi HJ, Kim I, Lee HJ, et al. Clinical characteristics of neonatal cholestasis in a tertiary hospital and the development of a novel prediction model for mortality. EBioMedicine 2022;77:103890.
60. Shalaby MA, Sawan ZA, Nawawi E, et al. Incidence, risk factors, and outcome of neonatal acute kidney injury: a prospective cohort study. Pediatr Nephrol

2018;33:1617-24.
61. Vardhelli V, Murki S, Tandur B, et al. Comparison of CRIB-II with SNAPPE-II for predicting survival and morbidities before hospital discharge in neonates with gestation ≤ 32 weeks: a prospective multicentric observational study. Eur J Pediatr 2022;181:2831-8.
62. Lee SK, Zupancic JA, Pendray M, et al. Transport risk index of physiologic stability: a practical system for assessing infant transport care. J Pediatr 2001;139:220-6.
63. Lee SK, Aziz K, Dunn M, et al. Transport Risk Index of Physiologic Stability, version II (TRIPS-II): a simple and practical neonatal illness severity score. Am J Perinatol 2013;30:395-400.
64. Richardson DK, Gray JE, McCormick MC, et al. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. Pediatrics 1993;91:617-23.
65. McLeod JS, Menon A, Matusko N, et al. Comparing mortality risk models in VLBW and preterm infants: systematic review and meta-analysis. J Perinatol 2020;40:695-703.
66. Gagliardi L, Cavazza A, Brunelli A, et al. Assessing mortality risk in very low birthweight infants: a comparison of CRIB, CRIB-II, and SNAPPE-II. Arch Dis Child Fetal Neonatal Ed 2004;89:F419-22.
67. Zhang L, Wu Y, Huang H, et al. Performance of PRISM III, PELOD-2, and P-MODS Scores in Two Pediatric Intensive Care Units in China. Front Pediatr 2021;9:626165.

Cite this article as: Liang Y, Zhao L, Huang J, Wu Y. A nomogram to predict 28-day mortality in neonates with sepsis: a retrospective study based on the MIMIC-III database. Transl Pediatr 2023;12(9):1690-1706. doi: 10.21037/tp-23-150

Appendix 1 Supplementary method

Missing data handling

This study is based on the MIMIC database. Although demographic data was mostly complete, laboratory indicators had several missing values, which, if eliminated, would lead to the loss of crucial information, thus invalidating our modeling approach. The missing data is classified into missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR) based on the cause of the missing values (68,69). To discover the nature of the data missing, the correlation matrix was performed to examine the correlation between missing values in the continuous variables (*Table S4*). To facilitate presentation, only significant variables such as weight, bicarbonate, calcium, serum sodium, creatinine, and base excess, were selected and their correlations were presented on the correlation matrix. The values ranged between −0.08 and 0.35, and a weak correlation could be observed between serum sodium and bicarbonate (0.35). Based on a comprehensive analysis, the missing data was deemed to be MCAR, and thus the missing values in the database were replaced with the method "norm. predict" in Multivariate Imputation by Chained Equation (MICE).

Details of the five assumptions in logistic regression

Assumption 1: appropriate outcome variable type

The statement appears to be correct in terms of meeting the first assumption of logistic regression, which requires a dichotomous outcome variable. The outcome variable in this study, death within 28 days, is binary, as it only has two possible outcomes: death or survival. Therefore, it satisfies the first assumption of logistic regression.

Assumption 2: linearity in the logit p (log OR)

The logistic regression model assumes a linear relationship between continuous variables and logit, which is an essential assumption. Two methods are commonly used to assess linearity: graphic visualization and the Box-Tidwell test.

Firstly, we performed restricted cubic spline analysis to explore potential nonlinear relationships between continuous variables and outcomes, a widely used method in this context (70-73). Our analysis did not reveal any significant nonlinearity relationship in the continuous variables, as shown in Figure 3 for both the lasso regression model and the binary logistic regression model.

Additionally, we performed the Box-Tidwell test to further validate our findings and ensure their robustness. As shown in *Figure 3*, all P-nonlinear >0.05, confirm that no significant nonlinearity relationship existed between the continuous variables and the outcome.

Assumption 3: multicollinearity

To diagnose multicollinearity, both the correlation coefficient and variance inflation factor (VIF) are useful metrics (29). First, Pearson and non-parametric Spearman correlation matrices are calculated to explore the possibility of multicollinearity in both continuous and bivariate analyses. Secondly, VIF values equal or greater than 5 indicate the presence of multicollinearity among variables (30). Importantly, neither the correlation matrix nor the VIF values revealed any significant multicollinearity between variables as shown in *Table 4* and *Figure S1*.

Assumption 4: independence of observations.

The assumption of independence in statistical analysis refers to the occurrence of positive events that are randomly distributed across different spaces, times, and populations (excluding the independent variables included in the model). In our study, the positive outcome is not influenced by the aforementioned factors. Therefore, we can assume that the independence hypothesis is approximately met.

Assumption 5: sample size

In this study, we employed one of the most recent techniques to calculate the sample size (27), and traditional methods that satisfy the criteria proposed by Peduzzi of event per variable (EPV) >10 suggest that 1,250 samples would be needed to achieve an EPV of 10 from the final selection of six variables and an event rate of 0.048. We failed to achieve sufficient statistical power. Acknowledging the limitations of the sample size, the validation of our conclusions will be required in future large-scale population studies.

References

68. Blazek K, van Zwieten A, Saglimbene V, et al. A practical guide to multiple imputation of missing data in nephrology. Kidney Int 2021;99:68-74.
69. Austin PC, White IR, Lee DS, et al. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. Can J Cardiol 2021;37:1322-31.
70. Tan Y, Fu Y, Yao H, et al. Relationship between phthalates exposures and hyperuricemia in U.S. general population,

https://dx.doi.org/10.21037/tp-23-150

a multi-cycle study of NHANES 2007-2016. Sci Total Environ 2023;859:160208.
71. Lee DH, Keum N, Hu FB, et al. Predicted lean body mass, fat mass, and all cause and cause specific mortality in men: prospective US cohort study. BMJ 2018;362:k2575.
72. Gao X, Wang J, Huang H, et al. Nomogram Model Based on Clinical Risk Factors and Heart Rate Variability for Predicting All-Cause Mortality in Stage 5 CKD Patients. Front Genet 2022;13:872920.
73. Weiser MR, Chou JF, Keshinro A, et al. Development and Assessment of a Clinical Calculator for Estimating the Likelihood of Recurrence and Survival Among Patients With Locally Advanced Rectal Cancer Treated With Chemotherapy, Radiotherapy, and Surgery. JAMA Netw Open 2021;4:e2133457.

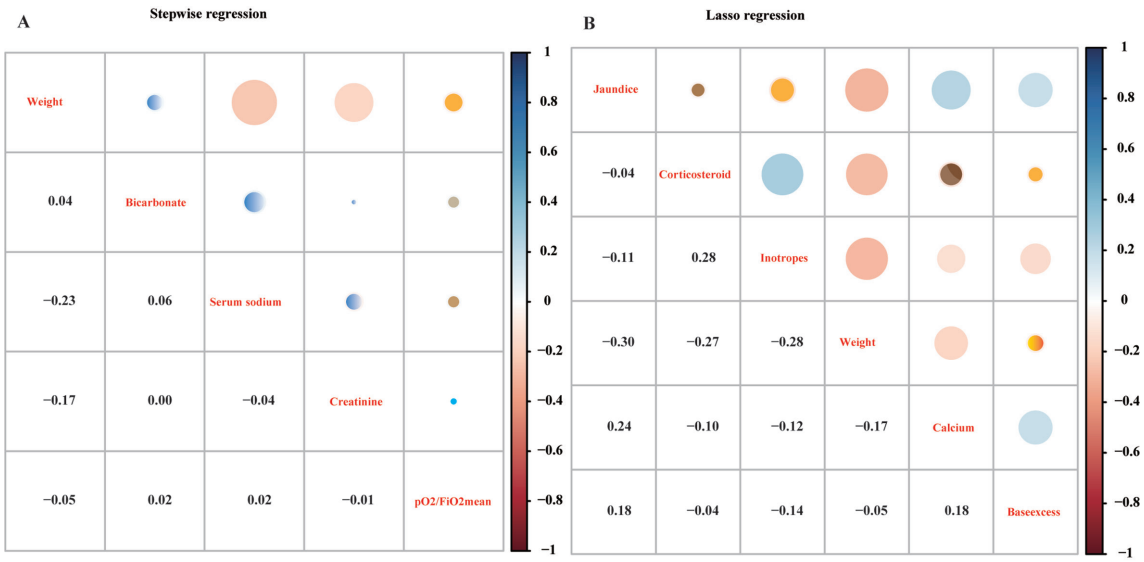

Figure S1 Pearson and nonparametric Spearman correlation matrices among variables in the models.

https://dx.doi.org/10.21037/tp-23-150

Table S1 Neonatal Sequential Organ Failure Assessment (nSOFA) Components and Scoring[†]

Component	nSOFA Scores				
Respiratory score	0	2	4	6	8
Criteria	Not intubated or intubated, SpO2/FiO2≥300	Intubated, SpO2/FiO2<300	Intubated, SpO2/FiO2<200	Intubated, SpO2/FiO2<150	Intubated, SpO2/FiO2<100
Cardiovascular score	0	1	2	3	4
Criteria[§]	No inotropes and no Systemic corticosteroid treatment	No inotropes and systemic corticosteroid treatment	1 inotrope and no systemic corticosteroid treatment	≥2 inotropes or 1 inotrope and systemic corticosteroid treatment	≥2 inotropes and systemic corticosteroid treatment
Hematologic score	0	1	2	3	NA
Criteria[¶]	Platelet count[‡] ≥150×10^9	Platelet count (100–149)×10^9	Platelet count <100×10^9	Platelet count <50×10^9	

[†], Score range, 0 (best) to 15 (worst). [‡], SI conversion factor: To convert platelet count to ×10^9/L, multiply by 1. [§], Medications considered as inotropic or vasoactive included dopamine, dobutamine, epinephrine, norepinephrine, vasopressin, and phenylephrine. [¶], Most recent platelet count available to the clinician. FiO2, fraction of inspiratory oxygen; SpO2, peripheral oximetric saturation; NA, not applicable.

Table S2 The correlation between missing values in all continuous variables

Variables	Weight	Bicarbonate	Calcium	Serum sodium	Creatinine	Base excess
Weight	1	0.3	0.04	0.3	–0.03	0.09
Bicarbonate	0.3	1	0.1	0.35	0.07	0.06
Calcium	0.04	0.1	1	0.1	0.12	–0.08
Serum sodium	0.3	0.35	0.1	1	0.07	0.06
Creatinine	–0.03	0.07	0.12	0.07	1	0.06
Base excess	0.09	0.06	–0.08	0.06	0.06	1

Table S3 Performance of the developed models and nSOFA

Models	AUC	Sensitivity	Specificity	Accuracy
Stepwise	0.784	0.998	0.037	0.952
Lasso	0.924	0.991	0.439	0.965
nSOFA	0.807	0.998	0.096	0.955

AUC, area under the receiver operating characteristic curve; Lasso, the least absolute shrinkage and selection operator; nSOFA, the neonatal sequential organ failure assessment score.

Table S4 The NRI and IDI estimate of the developed models and nSOFA.

Models	NRI		IDI	
	Estimate (95% CI), %	P value	Estimate (95% CI), %	P value
Stepwise regression and nSOFA	8.41 (–10.34, 27.15)	0.379	3.95 (–1.84, 9.74)	0.181
Lasso algorithm and stepwise regression	52.95 (35.70, 70.21)	<0.001	30.93 (21.94, 39.92)	<0.001
Lasso algorithm and nSOFA	53.64 (38.85, 68.43)	<0.001	26.98 (19.06, 34.89)	<0.001

nSOFA, the neonatal sequential organ failure assessment score; Lasso, the least absolute shrinkage and selection operator; NRI net reclassification improvement; IDI integrated discrimination improvement.

https://dx.doi.org/10.21037/tp-23-150