Predicting Mortality or Intestinal Failure in Infants with Surgical Necrotizing Enterocolitis

Darshna Bhatt, MHA, DO[1], Curtis Travers, MPH[2], Ravi M. Patel, MD, MSc[1], Julia Shinnick, MD[3], Kelly Arps, MD[3], Sarah Keene, MD[1], and Mehul V. Raval, MD, MS[3]

Objective To compare existing outcome prediction models and create a novel model to predict death or intestinal failure (IF) in infants with surgical necrotizing enterocolitis (NEC).

Study design A retrospective, observational cohort study conducted in a 2-campus health system in Atlanta, Georgia, from September 2009 to May 2015. Participants included all infants ≤37 weeks of gestation with surgical NEC. Logistic regression was used to model the probability of death or IF, as a composite outcome, using preoperative variables defined by specifications from 3 existing prediction models: American College of Surgeons National Surgical Quality Improvement Program Pediatric, Score for Neonatal Acute Physiology Perinatal Extension, and Vermont Oxford Risk Adjustment Tool. A novel preoperative hybrid prediction model was also derived and validated against a patient cohort from a separate campus.

Results Among 147 patients with surgical NEC, discrimination in predicting death or IF was greatest with American College of Surgeons National Surgical Quality Improvement Program Pediatric (area under the receiver operating characteristic curve [AUC], 0.84; 95% CI, 0.77-0.91) when compared with the Score for Neonatal Acute Physiology Perinatal Extension II (AUC, 0.60; 95% CI, 0.48-0.72) and Vermont Oxford Risk Adjustment Tool (AUC, 0.74; 95% CI, 0.65-0.83). A hybrid model was developed using 4 preoperative variables: the 1-minute Apgar score, inotrope use, mean blood pressure, and sepsis. The hybrid model AUC was 0.85 (95% CI, 0.78-0.92) in the derivation cohort and 0.77 (95% CI, 0.66-0.86) in the validation cohort.

Conclusions Preoperative prediction of death or IF among infants with surgical NEC is possible using existing prediction tools and, to a greater extent, using a newly proposed 4-variable hybrid model. *(J Pediatr 2017;191:22-7).*

Despite advances in neonatal care, necrotizing enterocolitis (NEC) remains a leading contributor to neonatal morbidity and mortality, accounting for 10% of deaths in the neonatal intensive care unit.[1,2] Among extremely preterm infants who have NEC and undergo surgical intervention, the mortality rate is as high as 50%.[3,4] Patients with NEC who require surgical intervention and survive are at risk for significant morbidities, which include the development of short bowel syndrome and intestinal failure (IF).[4-6] Despite the high risk of adverse outcomes for infants with surgical NEC, there are no validated risk prediction models for use in the preoperative period to inform discussions with families or guide risk adjustment comparisons within and between centers.[7-10] The ideal prediction model should be simple to use and include a small set of inputs that are easily accessible preoperatively, while being appropriately validated and calibrated.[11]

Existing models that predict disease severity among infants with NEC include the Stanford NEC model and the NEC-totalis model. However, these models do not focus on outcomes in the patient with surgical NEC.[9] A number of validated risk prediction models exist to predict neonatal mortality, including those who undergo surgery.[12] Of these, commonly used models include the Score for Neonatal Acute Physiology Perinatal Extension (SNAPPE-II), the Vermont Oxford Risk Adjustment Tool (VON-RA), and the American College of Surgeons National Surgical Quality Improvement Program Pediatric (NSQIP-P). The SNAPPE-II model was designed to predict death within the first 12 hours of life for newborn infants. The VON-RA model primarily functions to guide risk-adjusted comparisons of morbidity and mortality in very low birthweight infants.[10,13] The goal of the NSQIP-P

AUC	Area under the receiver operating characteristic curve
H-L	Hosmer-Lemeshow
IF	Intestinal failure
NEC	Necrotizing enterocolitis
NSQIP-P	American College of Surgeons National Surgical Quality Improvement Program Pediatric
SIP	Spontaneous intestinal perforation
SNAPPE-II	Score for Neonatal Acute Physiology Perinatal Extension Version II
VON-RA	Vermont Oxford Risk Adjustment Tool

From the [1]Division of Neonatology, Department of Pediatrics, Emory University School of Medicine, Children's Healthcare of Atlanta; [2]Biostatisitcal Core, Department of Pediatrics, Emory University School of Medicine; and [3]Division of Pediatric Surgery, Department of Surgery, Emory University School of Medicine, Children's Healthcare of Atlanta, Atlanta, GA

R.P receives support from the National Institutes of Health (NIH) (K23 HL128942). M.R. is supported by the Emory University Department of Surgery and the Emory Children's Pediatric Research Alliance. The other authors declare no conflicts of interest.

Portions of this study were presented at the Pediatric Academic Societies in Baltimore, Maryland, April 30-May 3, 2016; the American Academy of Pediatrics National Conference & Exhibition in Chicago, Illinois, September 16-19, 2017; and the Southern Society for Pediatric Research Annual Meeting in New Orleans, Louisiana, February 18-20, 2017.

https://doi.org10.1016/j.jpeds.2017.08.046

risk calculator model is to use preoperative variables to estimate the probability of adverse postoperative events in the pediatric patient population, including neonates.[6] The performance of these currently available tools in predicting mortality or IF among infants with surgical NEC is unknown.[14-16] The primary purpose of this study was to evaluate and compare these existing tools in the prediction of death or IF in patients with surgical NEC. A secondary aim was to derive and validate a novel hybrid model to predict death or IF using preoperative variables.

Methods

A retrospective, observational cohort study was conducted at 2 level IV neonatal intensive care units, based on the American Academy of Pediatrics designation,[17] from September 1, 2009, to May 31, 2015. Both free-standing children's hospitals were part of the same healthcare system (Children's Healthcare of Atlanta), but staffed by different neonatal and surgical practices. Infants with a gestational of age less than or equal to 37 weeks and a diagnosis of NEC receiving surgical intervention were included. Surgical intervention was defined as receipt of an exploratory laparotomy or primary percutaneous drain placement. Patients with a diagnosis of spontaneous intestinal perforation (SIP) or preexisting congenital intestinal anomalies were excluded. Determination of SIP was based on a review of the surgeon's operative report and clinical presentation (pneumoperitoneum in the first 7 days of life without radiographic evidence of NEC). We defined congenital intestinal anomalies as omphalocele, gastroschisis, small and large intestinal atresia, other intestinal obstructions present at birth, and malrotation. The primary outcome measure was death or IF. Death was defined as all-cause in-hospital mortality. IF was defined as the failure to achieve full enteral feeds at 90 days postoperatively, based on prior studies.[18,19] A cohort of infants from 1 hospital (derivation cohort) was used to assess the performance of existing prediction models and derive a new hybrid model, which was validated using a separate cohort from a second hospital (validation cohort).

Variables from the VON-RA, SNAPPE-II, and NSQIP-P models, along with other baseline, preoperative variables were extracted from electronic medical records and verified by individual data extractors. Where applicable, preexisting definitions from SNAPPE-II, NSQIP-P, and VON-RA were used. A definition manual was created, which included prespecified variable definitions (**Table III**). All variables collected were measured before surgery (closest to the time of surgery) and physiologic parameters were obtained within 3 hours before surgery. Variables from the SNAPPE-II, VON-RA, and NSQIP-P models are listed in **Table VII**. All variables for these models were extracted according to established definitions.[14,20,21]

Statistical Analyses

Counts and percentages were calculated for categorical variables and compared with the outcomes of death or IF, and death only using χ^2 tests or Fisher exact tests. Continuous variables were summarized with medians and IQRs and compared using Wilcoxon rank-sum tests. Multivariable logistic regression models were then fitted to model the probability of death or IF as a composite outcome for each of the models described. SNAPPE-II was scored following the defined methods, whereas VON-RA and NSQIP-P were modeled using their respective individual variables (**Tables III**, **IV**, and **V**; available at www.jpeds.com). To create a hybrid model, candidate variables were selected from variables with unadjusted *P* values of less than .1 (**Tables VI**, **VII**, and **VIII**; available at www.jpeds.com). Automated backward stepwise selection was used to select variables included in the final model to maximize the area under the receiver operating characteristic curve (AUC) using SAS version 9.4 (SAS Institute Inc, Cary, North Carolina). To protect against model overfitting, variables were removed from the model until the model contained 5 or fewer variables and further removal of variables would result in reductions in an AUC of less than 0.2. The least significant variable was dropped from the model in this process. Other selection strategies, including forward selection, were used to verify that the same variables were selected. Final model fit was assessed using Hosmer-Lemeshow (H-L) tests and graphically represented using calibration plots that plotted deciles of predicted probabilities of outcome with the corresponding observed risk of outcome. The final coefficients from the hybrid model, including the intercept, were then used to evaluate discrimination and calibration of the model using a validation cohort of patients with surgical NEC from a level IV neonatal intensive care unit with different practices and healthcare staff. The predicted probability of outcome was calculated using the model coefficients from the derivation cohort and the resulting probabilities were compared with observed results using calibration plots, as noted. The AUC for the hybrid model was compared with the other models using χ^2 tests. Similar methods were used to determine variable selection for a hybrid model to predict the outcome of mortality alone, which was derived and validated in a similar manner using the 2 separate cohorts.

Results

A total of 147 infants met the selection criteria for the derivation cohort and 76 infants met selection criteria for the validation cohort (**Figure 1**). In the derivation cohort, the median gestational age and birth weight were 27.1 weeks (IQR, 25.6-30.1) and 940 g (IQR, 740-1361), respectively (**Table I**). In addition, 60% of infants were male and 83% were singleton births. Overall patient characteristics did not differ between the derivation and validation cohorts, except for lower use of a primary percutaneous drain in the derivation cohort (18% vs 8%; $P = .02$). The incidence of mortality or IF was 64% and 70% in the derivation and validation cohorts, respectively (**Table I**). The NSQIP-P model demonstrated better discrimination of infants with death or IF, when compared with the SNAPPE-II and VON-RA models, AUC 0.84 (95% CI, 0.78-0.91) vs AUC 0.60 (95% CI, 0.48-0.72) and AUC 0.74 (95% CI, 0.66-0.84) (**Table II**).

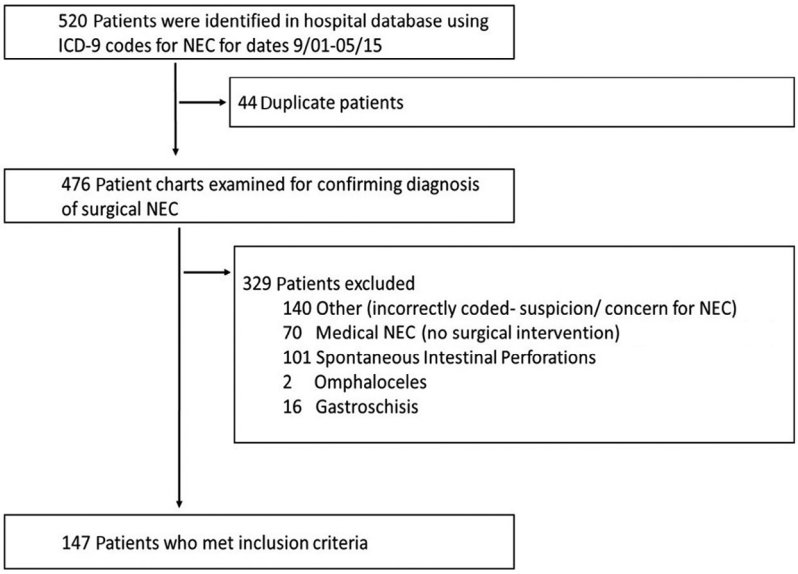

Figure 1. Selection of sample for model derivation. ICD-9, *International Classification of Diseases, 9th edition.*

Hybrid Model

After existing model evaluation, the following 4 variables were selected to predict the composite outcome of death or IF in a new hybrid model: 1-minute Apgar score, inotrope use, mean blood pressure, and sepsis (**Table VII**). The hybrid model had an AUC of 0.85 (95% CI, 0.78-0.92) for predicting mortality or IF in the derivation cohort and 0.78 (95% CI, 0.66-0.89) in the validation cohort. In comparison with the SNAPPE-II, the AUC was significantly greater using the hybrid model (**Table II**). The hybrid model showed modest improvements over NSQIP-P and VON-RA, using fewer variables. There was good fit in both derivation (H-L $P = .754$) and validation cohorts (H-L $P = .240$) for the hybrid model. **Figure 2** shows the observed vs predicted probabilities for death or IF in the derivation and validation cohorts, respectively. Of note, the hybrid model demonstrated sepsis to be associated with a lower risk of death or IF. This could be explained by a shorter duration of time for infants who died to have a blood culture result as positive at the time of preoperative ascertainment in this study. A prior study noted that the median time from onset of NEC and death was 1 day.[22]

For our secondary hybrid model, the following 5 variables were identified to predict mortality alone: preoperative inotrope use, mean blood pressure, PaO_2/fractional inspired oxygen ratio, hypothermia, and serum pH (**Table VII**). The hybrid model demonstrated good discrimination in predicting mortality (AUC, 0.81; 95% CI, 0.74-0.89) and fit (H-L $P = .481$). The model retained good discrimination (AUC, 0.88; 95% CI, 0.80-0.96) and fit (H-L $P = .384$) in the validation cohort. The hybrid model for mortality had nonstatistically significant improved discrimination compared with SNAPPE-II and VON-RA, and similar discrimination in comparison to NSQIP-P (**Table II**). Model intercept and coefficients are provided in **Table IX** (available at www.jpeds.com).

Table I. Infant characteristics

Patient characteristics	Derivation cohort (n = 147)	Validation cohort (n = 76)	*P* values
Gestational age, median weeks (IQR)	27.1 (25.6-30.1)	27.0 (25.8-29.9)	.55
Birthweight, median g (IQR)	940 (740-1361)	880 (720-1260)	.64
Male	88 (60)	41 (54)	.39
Female	59 (40)	35 (46)	
Antenatal steroids (≥1 doses)	95 (67)	44 (58)	.19
Initial surgical approach			
Percutaneous drain	12 (8)	14 (18)	.02
Exploratory laparotomy	135 (92)	62 (82)	
Mortality or IF (primary outcome)	94 (64)	53 (70)	.39
Mortality	56 (38)	27 (36)	.71
IF	52 (35)	26 (34)	.86

Values are n (%) unless otherwise noted.

Discussion

We found that the existing neonatal outcome prediction models, including the NSQIP-P, SNAPPE-II, and VON-RA, were able to predict mortality as well as the composite outcome of mortality or IF among infants with NEC. The NSQIP-P model demonstrated the best discrimination among the 3 models. Differences in the performance of the 3 models may be due to the predominant use of baseline characteristics in the SNAPPE-II and VON-RA models, which may be less useful in predicting outcomes that often occur weeks to months after birth such as in the setting of NEC. Although the NSQIP-P model demonstrated good ability to predict outcomes, the model requires more than 10 variables and, therefore, may not be pragmatic in use as a preoperative prediction or risk adjustment tool. The strength of the NSQIP-P is that is it uses

Table II. Performance of prediction models as assessed by the AUC for both death or IF and death alone

	Outcome: death or IF AUC‡‡ (95% CI)	*P* value* (comparison to hybrid)	Outcome: death AUC (95% CI)	*P* value* (comparison with hybrid)
Derivation cohort (n = 147)				
SNAPPE-II† overall score (n = 69)	0.60 (0.48-0.72)	.001	0.68 (0.53-0.82)	.045
VON-RA‡ (n = 137)	0.74 (0.66-0.84)	.051	0.67 (0.58-0.77)	.026
NSQIP-P§ (n = 135)	0.84 (0.77-0.91)	.85	0.82 (0.74-0.89)	.94
Hybrid (n = 135)	0.85 (0.78-0.92)¶	—	0.81 (0.74-0.89)**	—
Validation cohort (n = 76)				
Hybrid (n = 76)	0.77 (0.66-0.89)	.25	0.88 (0.81-0.96)	.24

*Significance assessed at $P < .0125$ after Bonferroni adjustment for multiple comparisons (0.05/4 comparisons).
†SNAPPE-II, n = 69 owing to missing data.
‡VON-RA, n = 137 owing to missing elements in patient charts.
§NSQIP-P, n = 135 owing to missing elements in patient charts.
¶Model includes Apgar score at 1 minute, inotropic use, mean blood pressure, and sepsis, n = 135 owing to missing elements in patient charts.
**Model includes inotropic use, mean blood pressure, pO_2/fractional inspired oxygen ratio, hypothermic, serum pH, n = 135 owing to missing elements in patient charts.

physiologic variables that are temporally relevant to the surgical risk period. In comparison with the NSQIP-P model, the newly derived hybrid model was designed to be parsimonious and contains 4 variables that should be readily available before surgery, thus increasing the model's simplicity in predicting the probability of death or IF.

After additional external validation, this novel prediction model has the potential to assist clinicians in prognostication and consultation with families preoperatively. Given the high incidence of death or IF among infants with surgical NEC, the ability to individualize the risk estimation of death or an adverse outcome, such as IF, could assist clinicians in

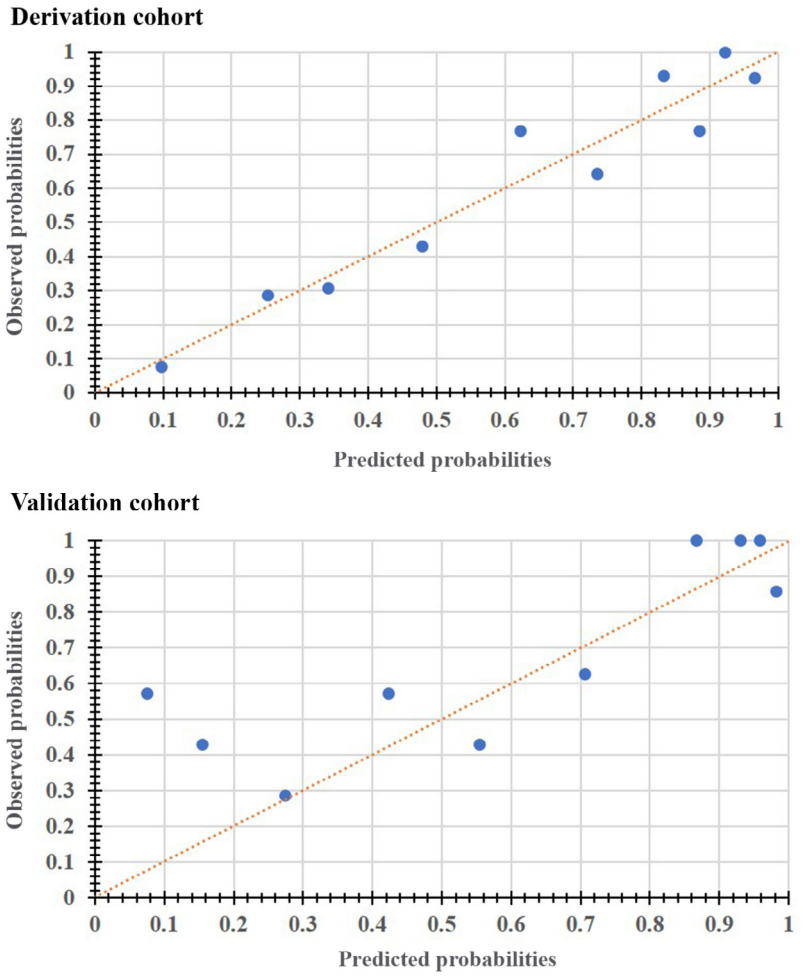

Figure 2. Observed vs predicted probabilities for death or IF.

discussions with families in preparing them for outcomes after surgery. The objective of the hybrid model is to provide the clinician additional prognostic information as they make judgements based on probabilities and the overall clinical scenario in counseling parents about both surgical risk and overall outcomes.

Furthermore, our prediction model has the potential to function as a tool for hospital-level outcomes comparisons among centers performing surgery for infants with NEC. Our hybrid model provides a tool for referral centers to estimate and adjust for the preoperative probability of an adverse outcome. This can then be used to provide less biased risk-adjusted comparisons of surgical outcomes among centers and help centers to understand how their surgical outcomes among infants with NEC compare with other centers.

Other prediction tools are currently available and used to guide neonatal care. One example is the National Institute of Child Health and Human Development neonatal outcome estimator, which uses 5 antenatal variables to predict the probability of survival without profound neurodevelopmental impairment at a corrected age of 18-22 months.[23] Of note, the AUC of this model (0.75; 95% CI, 0.74-0.77) was similar to the AUC of our hybrid model in the validation cohort for the composite outcome of death or IF (0.78; 95% CI, 0.66-0.89). Another tool is the Stanford NEC scoring system, which has been used to guide staging of infants at initial presentation of NEC.[24] The model uses variables, such as the presence of metabolic acidosis, portal venous gas, and abdominal wall discoloration, to predict the severity of NEC at the timing of onset of clinical symptoms. In contrast with our model, the Stanford NEC model does not estimate the probability of a poor outcome among infants undergoing surgery. More recently, Sho et al[9] reported a NEC-totalis prediction model, demonstrating that laboratory variables and clinical patient characteristics can predict NEC totalis. One of the limitations of this model is the large number of variables used, including the use of laboratory data that may not be available in the immediate preoperative period. Similar to our study, De Souza et al[25] examined both preoperative and postoperative predictive factors in identifying mortality in patients with surgical NEC in Brazil who underwent exploratory laparotomy. Patients with suspected SIP were excluded in this study, and both preoperative medical records and postoperative surgical reports were examined. The study found that intrauterine growth restriction and diffuse bowel involvement were the 2 factors predictive of mortality. This model differs from our model in that it does not solely use preoperative data, and clinicians may be limited in the prolonged time for observation and data collection necessary for the use of this model.

One strength of this study is the identification and exclusion of patients with SIP, who are typically not at risk of IF and have a lower mortality rate than infants with NEC.[24,26] This consideration is important, because up to 20% of patients in national surgical NEC datasets may actually have SIP and thus reported outcomes may be biased.[24] In addition, our hybrid model was validated in a separate cohort of patients who were cared for by a different group of neonatologists and pediatric surgeons than in the derivation cohort. Finally, we had a relatively large sample of infants with surgical NEC, which allowed for a sufficient sample to detect potential predictors of death or IF among infants with surgical NEC.

This study has several limitations, including the retrospective collection of data. Our comparison of models may have been impacted by missing data, although the degree of missing data for each of the models may be reflective of the real-world ability to ascertain variables included in these models before surgery. Therefore, we did not perform imputation for missing data. In particular, urine output measurements before surgery may be difficult to ascertain and limit the use of the SNAPPE-II model, which had the lowest discrimination of death or IF in our assessments. Prospective, real-time collection of variables immediately before the decision to proceed with surgery may be more informative. Therefore, the performance of our model using real-time or prospectively collected data immediately before the decision to proceed with surgery needs additional study. In addition, our study did not evaluate other major morbidities among infants with surgical NEC, such as long-term growth and neurodevelopmental outcomes. This is important to note, because the data suggest that infants with IF are among the highest risk for these adverse outcomes.[5,18,27] Furthermore, our model was validated against a cohort of patients from a hospital within the same healthcare system as the derivation cohort. Therefore, further external validation of our hybrid model in a geographically distinct group of patients is necessary before this model can be used clinically to provide prognostic estimates to families. The last limitation to acknowledge is the issue of self-fulling prophecy of risk prediction modeling.[28] It is possible that infants at the study centers considered to have a high probability of a poor outcome had death after withdrawal of care and that approaches to care may differ at other centers. This factor further emphasizes the importance of external validation of this model at other centers.

Prediction of death or IF among infants with surgical NEC is possible using existing prediction tools and, to a greater extent, using a newly proposed hybrid model. ■

Submitted for publication Apr 28, 2017; last revision received Jul 18, 2017; accepted Aug 17, 2017

Reprint requests: Darshna Bhatt, DO, MHA, Department of Pediatrics, Division of Neonatal-Perinatal Medicine, Children's Hospital of Richmond at Virginia Commonwealth University. E-mail: darshna.bhatt@vcuhealth.org; darshnabhatt@gmail.com

References

1. Lin PW, Stoll BJ. Necrotising enterocolitis. Lancet 2006;368:1271-83.
2. Neu JWW. Necrotizing enterocolitis. N Engl J Med 2011;364:255-64.
3. Fitzgibbons SC, Ching Y, Yu D, Carpenter J, Kenny M, Weldon C, et al. Mortality of necrotizing enterocolitis expressed by birth weight categories. J Pediatr Surg 2009;44:1072-6.
4. Dominguez KM, Moss RL. Necrotizing enterocolitis. Clin Perinatol 2012;39:387-401.

5. Raval MV, Moss RL. Current concepts in the surgical approach to necrotizing enterocolitis. Pathophysiology 2014;21:105-10.
6. Stey AM, Kenney BD, Moss RL, Hall BL, Berman L, Cohen ME, et al. A risk calculator predicting postoperative adverse events in neonates undergoing major abdominal or thoracic surgery. J Pediatr Surg 2015;50:987-91.
7. Rhee D, Salazar JH, Zhang Y, Yang J, Yang J, Papandria D, et al. A novel multispecialty surgical risk score for children. Pediatrics 2013;131:e829-36.
8. Dillon P, Hammermeister K, Morrato E, Kempe A, Oldham K, Moss L, et al. Developing a NSQIP module to measure outcomes in children's surgical care: opportunity and challenge. Semin Pediatr Surg 2008;17:131-40.
9. Sho S, Neal MD, Sperry J, Hackam DJ. A novel scoring system to predict the development of necrotizing enterocolitis totalis in premature infants. J Pediatr Surg 2014;49:1053-6.
10. Zupancic JA, Richardson DK, Horbar JD, Carpenter JH, Lee SK, Escobar GJ. Revalidation of the score for neonatal acute physiology in the Vermont Oxford Network. Pediatrics 2007;119:e156-63.
11. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
12. Medlock S, Ravelli ACJ, Tamminga P, Mol BWM, Abu-Hanna A. Prediction of mortality in very premature infants: a systematic review of prediction models. PLoS ONE 2011;6:e23441.
13. Manukyan N, Eppstein MJ, Horbar JD, Leahy KA, Kenny MJ, Mukherjee S, et al. Exploratory analysis in time-varying data sets: a healthcare network application. Int J Adv Comput Sci 2013;3:322-9.
14. Richardson DK, Corcoran JD, Escobar GJ, Lee SK. SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores. J Pediatr 2001;138:92-100.
15. Stey A, Barnert ES, Tseng C-H, Keeler E, Needleman J, Leng M, et al. Outcomes and costs of surgical treatments of necrotizing enterocolitis. Pediatrics 2015;135:8.
16. Zupancic JAF, Richardson DK, Horbar JD, Lee SK, Escobar GJ. Vermont oxford network SNAP pilot project participants. Revalidation of the score for neonatal acute physiology in the Vermont Oxford Network. Pediatrics 2007;119:e156-63.
17. American Academy of Pediatrics Committee on Fetus and Newborn. Levels of neonatal care. Pediatrics 2012;130:587.
18. Kelleher J, Mallick H, Soltau TD, Harmon CM, Dimmitt RA. Mortality and intestinal failure in surgical necrotizing enterocolitis. J Pediatr Surg 2013;48:568-72.
19. Duro D, Kalish LA, Johnston P, Jaksic T, McCarthy M, Martin C, et al. Risk factors for intestinal failure in infants with necrotizing enterocolitis: a Glaser Pediatric Research Network Study. J Pediatr 2010;157:203-8, e1.
20. Vermont Oxford Network. In: Vermont Oxford Network, ed. Manual of operations (part 2). release 19.0 ed. 2014. p. 1-97 https://public.vtoxford.org/wp-content/uploads/2014/11/Manual_of_Operations_Part2_v19.pdf. Vermont Oxford Network.
21. American College of Surgeons. National Surgical Quality Improvement Program. Chapter 4 ACS NSQIP Pediatric Variables and Definitions. In: ACS NSQIP Pediatric Operations Manual. Chicago, IL: American College of Surgeons; 2016.
22. Clark RH, Gordon P, Walker WM, Laughon M, Smith PB, Spitzer AR. Characteristics of patients who die of necrotizing enterocolitis. J Perinatol 2012;32:199-204.
23. Tyson JE, Parikh NA, Langer J, Green C, Higgins RD. Intensive care for extreme prematurity — moving beyond gestational age. N Engl J Med 2008;358:1672-81.
24. Ji J, Ling XB, Zhao Y, Hu Z, Zheng X, Xu Z, et al. A data-driven algorithm integrating clinical and laboratory features for the diagnosis and prognosis of necrotizing enterocolitis. PLoS ONE 2014;9:e89860.
25. De Souza JCK, Da Motta UIC, Ketzer CR. Prognostic factors of mortality in newborns with necrotizing enterocolitis submitted to exploratory laparotomy. J Pediatr Surg 2001;36:482-6.
26. Wadhawan R, Oh W, Hintz SR, Blakely ML, Das A, Bell EF, et al. Neurodevelopmental outcomes of extremely low birth weight infants with spontaneous intestinal perforation or surgical necrotizing enterocolitis. J Perinatol 2013;34:64.
27. Patel RM. Short- and long-term outcomes for extremely preterm infants. Am J Perinatol 2016;33:318-28.
28. Rysavy MA, Tyson JE. The problem and promise of prognosis research. JAMA Pediatr 2016;170:411-2.

Table III. Bivariate association between characteristics and mortality or IF for SNAPPE-II

Patient Characteristics†	Overall (n = 147)	Mortality/no feed (n = 94)	Non mortality/yes feed (n = 53)	OR (95% CI)	*P* value
SNAPPE-II†					
Mean blood pressure (per 1-unit increase) (n = 143)	43.2 (11.8)	40.2 (11.7)	48.4 (10.2)	0.94 (0.91-0.97)	**<.001***
Lowest temperature (24 hours prior) (per 1°C increase) (n = 146)	36.8 (36.5-37.0)	36.7 (36.3-37.0)	36.9 (36.7-37.0)	0.49 (0.25-0.98)	.044
pO_2/FiO_2 ratio (n = 144) (per 10-unit increase)	49.0 (1.6-153.9)	32.5 (1.1-122.1)	105.0 (4.1-198.5)	0.95 (0.92-0.99)	**.010***
Serum pH (n = 144) (per 0.1-unit increase)	7.26 (7.13-7.32)	7.20 (7.08-7.31)	7.29 (7.19-7.37)	0.65 (0.51-0.84)	**.001***
Multiple seizures (n = 144)	5 (3.5)	4 (4.3)	1 (2.0)	2.25 (0.24-20.66)	.474
Urine output (n = 72)	2.68 (1.16-3.83)	2.46 (1.04-3.70)	2.96 (1.16-3.96)	0.95 (0.74-1.21)	.897
Overall score (per 1-point increase)	0 (0-5)	0 (0-19)	0 (0-5)	1.07 (1.00-1.14)	.055

Values are median (IQR) or n (%) unless otherwise noted.
*Bolded *P* values were considered candidate variables for hybrid model for mortality or IF.
†ORs are reflective of a 1-point increase unless otherwise stated in the table.

Table IV. Bivariate association between characteristics and mortality or IF for NSQIP-P

Patient characteristics	Overall (n = 147)	Mortality/no feed (n = 94)	Non mortality/yes feed (n = 53)	OR (95% CI)	*P* value
NSQIP-P					
Case type					
Emergent	147 (100.0)			—	
Elective	0 (0.0)				
Urgent	0 (0.0)				
Premature (≤36 weeks of gestational age)	138 (93.9)	91 (96.8)	47 (88.7)	3.87 (0.93-16.17)	**.071**
Small for gestational age (n = 145)	10 (6.9)	6 (6.5)	4 (7.6)	0.86 (0.23-3.18)	.815
Birth weight (per 100 g)	940 (740-1361)	848 (690-1263)	1170 (811-1613)	0.93 (0.88-0.98)	**.007**
Nutritional supplement (n = 145)	70 (48.3)	43 (46.2)	27 (51.9)	0.80 (0.40-1.57)	.511
Dialysis	0 (0.0)	0 (0.0)	0 (0.0)	—	>.99
Cardiac risk factors (n = 145)					
Severe	2 (1.4)	2 (2.2)	0 (0.0)		—
Major	45 (31.0)	24 (25.8)	21 (40.4)	0.46 (0.21-1.01)	.052
Minor	28 (19.3)	16 (17.2)	12 (23.1)	0.50 (0.20-1.24)	.134
None	70 (48.3)	51 (54.8)	19 (36.5)	Reference	—
Apgar at 1 min	5 (3-7)	5 (3-7)	7 (4-8)	0.82 (0.71-0.95)	**.008***
Apgar at 5 min	8 (6-9)	8 (7-9)	9 (6-9)	0.94 (0.77-1.14)	.531
Inotropic use (n = 145)	87 (60.0)	70 (75.3)	17 (32.7)	6.27 (2.97-13.22)	**<.001***
Blood transfusion	92 (62.6)	59 (62.8)	33 (62.3)	1.02 (0.51-2.05)	.952
Hepatobiliary disease	20 (13.6)	14 (14.9)	6 (11.3)	1.37 (0.49-3.81)	.545
Hereditary bleeding disease	0 (0.0)	0 (0.0)	0 (0.0)	—	>.99
Sepsis	36 (24.5)	18 (19.2)	18 (34.0)	0.46 (0.21-0.99)	**.047***
Intracranial hemorrhage	1 (0.7)	1 (1.1)	0 (0.0)	—	>.99

Values are median (IQR) or n (%) unless otherwise noted.
*Bolded *P* values were considered candidate variables for hybrid model for mortality or IF.

Table V. Bivariate association between characteristics and mortality or IF for the Vermont Oxford Risk-Adjusted Model

Patient characteristics*	Overall (n = 147)	Mortality/no feed (n = 94)	Non mortality/yes feed (n = 53)	OR (95% CI)	*P* value†
VON*					
Apgar at 1 min	5 (3-7)	5 (3-7)	7 (4-8)	0.82 (0.71-0.95)	**.008**
Apgar at 5 min	8 (6-9)	8 (7-9)	9 (6-9)	0.94 (0.77-1.14)	.531
Gestational age‡	27 (25-30)	26 (25-29)	29 (26-32)	0.86 (0.78-0.94)	**.001**
Gestational age with days§	27.1 (25.6-30.1)	26.9 (25.4-29.3)	29.1 (26.6-32.4)	0.86 (0.78-0.94)	**.001**
Birth defects	4 (2.7)	3 (3.2)	1 (1.9)	1.71 (0.17-16.89)	.645
Sex					
Male	88 (59.9)	54 (57.5)	34 (64.2)	0.75 (0.38-1.51)	.426
Female	59 (40.1)	40 (42.5)	19 (35.9)	Reference	
Singleton birth	122 (83.0)	77 (81.9)	45 (84.9)	0.81 (0.32-2.02)	.644
Delivery (n = 146)					
Caesarean	90 (61.6)	62 (66.7)	28 (52.8)	1.79 (0.90-3.56)	.100
Vaginal	56 (38.4)	31 (33.3)	25 (47.2)	Reference	

Values are median (IQR) or n (%) unless otherwise noted.
*ORs are reflective of a 1-point increase unless otherwise stated in the table.
†Bolded *P* values were considered candidate variables for hybrid model for mortality or IF.
‡Presented as completed weeks of gestation.
§Presented as exact gestational age, including weeks and days.

Table VI. Variable definitions

Variable names	Definition	Variable input (answer)
Inotropic use	Intravenous inotropic pharmacologic support required before time of surgery, within 24 hours prior to primary surgical intervention Dopamine (low-dose dopamine [<5 μg] still counts as yes) Milrinone Vasopressin Dobutamine Epinephrine Norepinephrine Isoproterenol Ephedrine Inamrinone	Yes or no
Mean blood pressure	Preoperative documentation of mean blood pressure within 3 hours before primary surgical intervention, record mean blood pressure as close to time of surgery as possible.	A numerical value in mm Hg
Lowest Temperature	Preoperative documentation of temperature within 3 hours before primary surgical intervention as close to time of surgery as possible.	A numerical value (in degrees Celsius)
Lowest serum pH	Preoperative documentation of serum pH, within 3 hours before primary surgical intervention as close to time of surgery as possible.	A numerical value without units
Serum partial pressure of oxygen (pO_2)	Preoperative documentation of serum pO_2, within 3 hours before primary surgical intervention as close to time of surgery as possible.	A numerical value in mm Hg
Fraction of inspired oxygen	Preoperative documentation of fractional inspired oxygen, within 3 hours before primary surgical intervention as close to time of surgery as possible.	A numerical value without units
Sepsis	Preoperative documentation of blood culture positive sepsis (within 7 days) before primary surgical intervention.	Yes or no
Full feeds at 90 days	All feeds via oral, orogastric, nasogastric, gastrostomy, nasojejunal, or jejunostomy tubes at 90 days after the primary surgical intervention.	Yes or no
Mortality	Death before discharge from the hospital.	Yes or no

Table VII. Model variables*

Hybrid models†	SNAPPE-II model[14]	NSQIP-P	VON-RA
Mortality/IF	**Mean blood pressure**	Preterm	Gestational age
Apgar score at 1 min	Temperature	Small for gestational age (<10th percentile)	Gestational age (squared)
Inotrope use	pO_2/FiO_2	Birthweight	Multiple gestation/singleton gestation
Mean blood pressure	**Lowest serum pH**	Mode of delivery	Outborn status
Sepsis	Multiple seizures	Case type	**Apgar score at 1 min**
Mortality alone	Urine output	Nutritional supplement	Apgar score at 5 min
Inotrope use		Dialysis	Sex
Mean blood pressure		Cardiac risk factors	Cesarean delivery
pO_2/fractional inspired oxygen ratio		Apgar score at 1 min	Presence of a congenital anomaly
Hypothermic		Apgar score at 5 min	
Lowest serum pH		**Inotrope use 24 hr preoperatively**	
		Blood transfusion 48 hr before operative procedure	
		Hepatobiliary disease	
		Hereditary bleeding disorder	
		Intracranial hemorrhage	
		Systemic sepsis	

FiO_2, fraction of inspired oxygen; *pO_2*: partial pressure of oxygen.
*Variables from existing models incorporated into hybrid model predicting death or IF are bolded.
†All variables for hybrid model measured preoperatively (within 3 hours of surgery) for measuring composite outcome of mortality or IF.

Table VIII. Hybrid model adjusted OR results for primary composite outcome (IF or death) and mortality alone*

	Derivation cohort (n = 147)		
Effects	**OR (95% CI)**	***P* value**	**AUC (95% CI)**
Results for IF or mortality			
Apgar at 1 min (per 1-point increase)	0.76 (0.63-0.92)	.004	0.850
Inotropic use (yes)	6.16 (2.52-15.04)	<.001	(0.782-0.918)
Mean blood pressure (per 1-unit increase)	0.92 (0.88-0.97)	<.001	
Sepsis (yes)	0.27 (0.10-0.73)	.010	
Results for mortality			
Inotropic use (yes)	2.28 (0.94-5.58)	.070	0.813
Mean blood pressure (per 1-unit increase)	0.97 (0.93-1.00)	.065	(0.740-0.886)
pO_2/FiO_2 ratio (per 10-unit increase)	0.96 (0.91-1.01)	.094	
Preoperative hypothermia (yes)	2.50 (0.90-6.94)	.079	
Serum pH (per 0.1-unit increase)	0.70 (0.53-0.93)	.014	

*ORs are reflective of 1-point or 1-unit increase unless otherwise stated in the table.

Table IX. Multivariable logistic regression equations for probability of mortality or IF, or mortality alone

Probability of mortality or IF

$$P_{mortality/IF} = \frac{e^{4.787-0.275(Apgar\ at\ 1\ min)+1.818(inotropic\ use)-0.079(mean\ BP)-1.325(sepsis)}}{e^{1+4.787-0.275(Apgar\ at\ 1\ min)+1.818(inotropic\ use)-0.079(mean\ BP)-1.325(sepsis)}}$$

Probability of mortality

$$P_{mortality} = \frac{e^{26.387+0.826(inotropic\ use)-0.035(mean\ BP)-0.0044\left(\frac{PO_2}{FiO_2}\right)+0.915(hypothermic)-3.567(serum\ pH)}}{e^{1+26.387+0.826(inotropic\ use)-0.035(mean\ BP)-0.0044\left(\frac{PO_2}{FiO_2}\right)+0.915(hypothermic)-3.567(serum\ pH)}}$$

BP, mean blood pressure; *Hypothermic*, <36.5°C.
Variable specification: See variable specifications in **Table VI**.