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**).