UNIVERSITY OF CINCINNATI

Date: May 18, 2006

I, Jareen Meinzen-Derr, hereby submit this work as part of the requirements for the degree of:
Doctorate of Philosophy

in:
Epidemiology

It is entitled:
A Prediction Model for Risk of Necrotizing Enterocolitis among Very Low Birth Weight Infants

This work and its defense approved by:

Chair: Paul Succop, Ph.D.
Kim Dietrich, Ph.D.
Edward Donovan, M.D.
Richard Hornung, Ph.D.
Ardythe L. Morrow, Ph.D.
A Prediction Model for Risk of Necrotizing Enterocolitis Among Very Low Birth Weight Infants

A dissertation submitted to the

Division of Research and Advanced Studies
Of the University of Cincinnati

In partial fulfillment of the requirements for the degree of

DOCTORATE OF PHILOSOPHY (Ph.D.)

In the Division of Epidemiology and Biostatistics
of the Department of Environmental Health
of the College of Medicine

2006

by

Jareen Meinzen-Derr

B.S., University of Alabama at Birmingham – 1994
M.P.H., University of Alabama at Birmingham, School of Public Health – 1997

Committee Chair: Paul Succop, Ph.D.
Abstract

Necrotizing enterocolitis (NEC) is a sudden-onset life-threatening disease that occurs primarily in premature infants, with an incidence as high as 13% among sub-groups and a case-fatality rate near 40%. Although individual risk factors have previously been associated with NEC, most studies were small and had insufficient power to investigate several risk factors simultaneously. The objectives of this study were to 1) create and validate a statistical model to predict the development of NEC among infants 401-1500 grams birth weight; 2) compare statistical methodologies available for this purpose (classification and regression tree [CART] vs. logistic regression); and 3) statistically define a temporal cluster of NEC. This study is the first to address these objectives for NEC.

A prediction model for NEC among very low birth weight (VLBW) infants had reasonable accuracy (area under the ROC curve of 0.69, 95% CI 0.66, 0.73) utilizing 6 factors: birth weight, race, patent ductus arteriosus, enteral feeding, parity, and maternal hypertension. Further investigation revealed birth weight specific risk factors for NEC, so additional models were developed stratified by birth weight. While results from the CART analysis were comparable to logistic regression, the CART analysis revealed additional variables that were important only in certain sub-groups.

Although sporadic NEC cases are most common, NEC outbreaks have also been reported. The current study developed a statistical methodology to identify NEC clusters, and evaluated clinical differences between clustered and sporadic cases of NEC. Between 1996 and 2004, the incidence rate (95% CI) of NEC was 6.9% (6.1, 7.9) and the incidence density rate was 1.29 per 1000 patient-days (1.11, 1.45). The incidence of NEC did not increase over time, nor was any seasonal variation detected. Seventeen temporal clusters of NEC were identified, but no clinically significant differences between clustered and sporadic cases were seen. Some of the infants had dysfunctional respiratory systems at birth, suggesting a pathway to NEC through intestinal ischemia. With further research on NEC clustering, the role of
environmental factors will be refined, allowing for more appropriate intervention. This study of NEC among VLBW infants highlights a fundamental need for continued epidemiologic research in this field.
Acknowledgements

I would first like to thank my members of my doctoral committee, Dr. Kim Dietrich, Dr. Edward Donovan, Dr. Richard Hornung, Dr. Ardythe L. Morrow, and Dr. Paul Succop, for providing guidance and advice through this entire process. I thank Dr. Succop, chairperson and advisor, for being an excellent teacher, facilitating in the development of my statistical skills and having faith in my abilities. I would like to express my gratitude to Drs. Morrow and Donovan, who have contributed to the development of my professional skills and provided constant encouragement. I am grateful to Dr. Hornung for allowing me to drop by and ask questions, and to Dr. Dietrich who has been extremely thoughtful throughout this process.

I would also like to express my gratitude to the National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN) for their continuing support and enthusiasm for my work as it relates to the current and similar projects. A special thanks is in order to the NRN Coordinators and Research Nurses, whose dedication to the NRN has made this epidemiologic work possible.

I would like to thank members in the Division of Neonatology for being excellent teachers of great patience. They have never hesitated to answer any and all questions I have had through this process, and have welcomed me into their Division with open arms. It is with their support and vast understanding of the importance of this work that has made this adventure both interesting and exciting.

Most importantly, I thank my family for constantly keeping me grounded throughout this process. My siblings, Jenara and Jaran, have made sure that I have had plenty of laughter during my frustrating moments. My parents, Dan and Jude, have made sure that I was always aware of what was truly important in life. Seth and Sebella, my wonderful nephew and niece, have and will continue to simply be a source of pure joy, particularly when I’ve needed it the most. And finally, I am especially grateful to Susan Wiley. Words cannot express how important her support and confidence in me has been throughout this experience. She always
made sure that I was eating when I was too busy to care, sleeping when I was too tired to think, and incorporating activities into my life that allowed me some fun. Her benevolent spirit and gentle soul has made me want to be a better person. Her presence has made this entire experience a little easier. Thank you.
# Table of Contents

TABLE OF CONTENTS ............................................................................................................................................ 1

ABBREVIATIONS .................................................................................................................................................. 3

LIST OF TABLES AND FIGURES ......................................................................................................................... 4

CHAPTER 1: INTRODUCTION .................................................................................................................................. 8

HYPOTHESES AND SPECIFIC AIMS .................................................................................................................... 9

CHAPTER 2: BACKGROUND ................................................................................................................................. 11

EPIDEMIOLOGY OF NECROTIZING ENTEROCOLITIS ........................................................................................... 11
  Complications and Death due to Necrotizing Enterocolitis ............................................................................. 12
  Pathogenesis and Risk Factors ......................................................................................................................... 13
  Birth weight and NEC ..................................................................................................................................... 14
  Feeding and NEC .......................................................................................................................................... 15

EPIDEMICS AND OUTBREAKS ............................................................................................................................ 17
  Clusters of Necrotizing Enterocolitis ................................................................................................................ 19

PREDICTION MODELS ....................................................................................................................................... 20
  Importance of Model Validation ..................................................................................................................... 21

RECURSIVE PARTITIONING ................................................................................................................................. 21

CHAPTER 3: METHODS ....................................................................................................................................... 24

DESCRIPTION OF STUDY POPULATION ........................................................................................................... 24

POWER CALCULATION ...................................................................................................................................... 24

SELECTION OF STUDY COHORT ........................................................................................................................ 25

DATA COLLECTION .......................................................................................................................................... 26

CASE DEFINITION .............................................................................................................................................. 26

OVERVIEW AND JUSTIFICATION OF STATISTICAL METHODS ........................................................................ 27
  Dichotomized and Calculated Variables ......................................................................................................... 27
  Variable Diagnostics ..................................................................................................................................... 28

LOGISTIC REGRESSION MODELING ................................................................................................................... 29
  Model Diagnostics and Goodness-of-Fit Statistics ......................................................................................... 30

RECURSIVE PARTITIONING ................................................................................................................................. 31

DETAILED ANALYSIS PLAN BY SPECIFIC AIM ............................................................................................... 33
  Model Development-Logistic Regression Modeling ....................................................................................... 34
  Evaluation of Model Performance .................................................................................................................. 35
  Comparison of the Final Model to a Simple Model ....................................................................................... 36

CHAPTER 4: RESULTS ....................................................................................................................................... 42

STUDY COHORT DESCRIPTION .......................................................................................................................... 42

PREDICTORS OF NEC ....................................................................................................................................... 43
  Unadjusted Results ......................................................................................................................................... 43
  Maternal/Prenatal Characteristics .................................................................................................................. 43
  Clinical Characteristics of Infant ................................................................................................................... 44

MODEL DEVELOPMENT ................................................................................................................................... 44

RESULTS OF LOGISTIC REGRESSION ANALYSIS ON THE DERIVATION SAMPLE ........................................... 46
  Combined Model Results ............................................................................................................................... 48
  Stratified models ........................................................................................................................................... 50
  Parity and NEC ............................................................................................................................................. 55
  Hypertension and NEC ................................................................................................................................ 56

EXTERNAL VALIDATION ..................................................................................................................................... 56

NEC OR DEATH ............................................................................................................................................... 59

RECURSIVE PARTITIONING ................................................................................................................................. 60
  Birth weight < 1000 grams ............................................................................................................................ 60
Abbreviations

AIC    Akaike’s Information Criterion
AUC    area under the curve
BTWT   birth weight
CI     confidence interval
d     days
DF     degrees of freedom
g     grams
GA     gestational age
H-L    Hosmer – Lemeshow
HMD    hyaline membrane disease
hrs    hours
IUGR   intrauterine growth retardation
NEC    necrotizing enterocolitis
OR     odds ratio
P     p-value
PDA    patent ductus arteriosus
PROM   prolonged rupture of membranes
RDS    respiratory distress syndrome
resp   respiratory
ROC    receiver operator characteristic
SD     Standard deviation
SE     standard error
SGA    small for gestational age
VLBW   very low birth weight
wk     week
List of Tables and Figures

Table 1. Bell’s Staging for Necrotizing Enterocolitis (NEC) ........................................................93

Figure 1: Pathogenesis of NEC consisting of 3 major factors....................................................94

Table 2: Proposed Risk Factors for NEC in Preterm Infants ......................................................95

Table 3: Published NEC Outbreaks ............................................................................................97

Figure 2: Example of generic classification and regression tree output ....................................102

Figure 3: Birth weight cut-off for stratified analysis. Cut-off based on the best sensitivity-
specificity trade off ....................................................................................................................103

Figure 4: Proportion of VLBW infants who are small for gestational age .................................104

Figure 5: Incidence and 95% Confidence Interval of Necrotizing Enterocolitis by Year ..........105

Figure 6: Incidence and 95% Confidence Interval of NEC by birth weight categories ..........106

Figure 7: Scatter plot and Regression Line of Age of NEC Onset and Birth Weight ..............107

Figure 8: Median and Interquartile Ranges of NEC by Birth Weight .....................................108

Figure 9: Mortality rates over time for infants with and without NEC ....................................109

Table 5: Comparison of Maternal/Pregnancy Characteristics between Derivation and Validation
Samples ....................................................................................................................................110

Table 6: Comparison of Clinical Characteristics between Derivation and Validation Samples 111

Table 7: Maternal/Pregnancy Characteristics of Derivation Sample .......................................113

Table 8: Clinical characteristics of VLBW infants for Derivation Sample ..............................114

Table 9: Logistic regression results from derivation sample ..................................................117

Table 10: Logistic regression results Derivation Sample with interaction term .....................118

Figure 10: Calibration plot for derivation model ......................................................................119

Table 11: Logistic regression results Validation Sample ........................................................120

Figure 11: Area under the ROC curve for Derivation and Validation Models .......................121

Table 12: Associations between clinical characteristics and the presence of NEC in the
derivation sample and the validation sample ............................................................................122
Figure 12: Calibration plots for both Derivation and Validation Models ........................................123
Table 13: Maternal/Pregnancy Characteristics; All infants ............................................................124
Table 14: Clinical characteristics of VLBW infants .......................................................................125
Table 15: Logistic regression results on total sample .....................................................................127
Table 16: Logistic regression results on total sample, no interaction between birth weight and parity .................................................................................................................................128
Figure 13: Predicted Probabilities from Logistic Regression Modeling for infants with and without NEC ..............................................................................................................................................129
Figure 14: Area under the ROC curve (95% confidence interval) for the final prediction model ..................................................................................................................................................130
Table 17: Comparison of Maternal/Pregnancy Characteristics by Birth Weight Category ............131
Table 18: Comparison of clinical characteristics by birth weight categories ....................................132
Table 19: Maternal characteristics of VLBW infants and NEC by birth weight category ............134
Table 20: Clinical characteristics of VLBW infants and NEC by birth weight category ...............135
Table 21: Stratified Multivariate Logistic Regression Results: Infants 401-999 g birth weight ..138
Table 22: Stratified Multivariate Logistic Regression Results: Infants 1000-1500g birth weight ..................................................................................................................................................139
Figure 15: Interaction between Birth Weight and Maternal Hypertension .....................................140
Figure 16: Area Under the Receiver Operator Curves (AUC) for 2 Multivariable Logistic Regression Models ...........................................................................................................................................141
Figure 17: Predicted Probabilities of infants with and without NEC .............................................142
Table 23: Logistic regression results on total sample with hypertension included .......................143
Table 24: Logistic regression results on total sample after adding hypertension and birth weight interaction term ..............................................................................................................................................144
Table 25: Comparisons of the area under the ROC curve and AIC candidate multivariable logistic models .............................................................................................................................................145
Table 26: Comparisons of the area under the ROC curve (AUC) for all univariable models compared to the final multivariable prediction model

Table 27: Maternal Characteristics According to Parity of Child

Table 28: Clinical characteristics of VLBW infants according to Parity of Child

Figure 18: Interaction between Birth Weight and Parity of infant

Table 29: Maternal characteristics according to pregnancy induced hypertension status

Table 30: Clinical characteristics of VLBW infants by pregnancy induced hypertension status

Table 31: Unadjusted associations between Maternal/Pregnancy Characteristics and NEC

Table 32: Unadjusted associations between clinical characteristics of VLBW infants and NEC

Figure 19: Calibration Plot for Multivariable Logistic Prediction Models: Internal Sample (N=3,816) and External Sample (n=32,770)

Table 33: Logistic regression results– Internal Sample and External Validation Sample

Table 34: Logistic regression results –Final Model External Validation Sample

Table 35: Logistic regression results infants 401-999 g– Internal Sample and External Validation Sample

Table 36: Logistic regression results infants 1000-1500 g; Internal Sample and External Validation Sample

Table 37: Comparison of Sensitivity and Specificity of Original and Externally Validated Models

Figure 20: Tree-based analysis of infants with NEC

Table 38: Comparison of Performance of CART and Logistic Regression

Figure 21: Number of NEC events and Incidence Rate over time

Figure 22: Overall Seasonal NEC Rate (1996-2004)

Figure 23: Incidence of NEC for Study Years 1996-2000 and 2001-2004

Table 39: Temporal Clusters identified within a Geographic Region
Chapter 1: Introduction

Necrotizing enterocolitis (NEC) is a sudden-onset life-threatening disease that occurs primarily in premature infants who have survived the critical first days of life, and occurs even among those who are considered to be “doing well”. The unpredictable nature of this devastating disease has made NEC the Russian roulette of prematurity. The incidence of this devastating disease has been reported to be as high as 13% among subgroups of premature infants, with a case fatality rate as high as 40%. Although the disease is well described and intensely studied, the cause or causes of NEC are not yet established. NEC in preterm infants is associated with a significant risk of death or long-term morbidities such as short-gut syndrome and cholestatic liver disease; a large proportion of survivors have poor long term growth and lifelong developmental handicaps. Both the incidence and severity of NEC have remained essentially constant in recent years, with few existing or proposed new treatments that show promise for treatment of NEC once it has developed.

At present it is difficult to predict which premature infants will develop NEC, and therefore, evaluation of preventive interventions is not efficient. If a potentially preventive intervention were given to all NICU infants, approximately 90% of the infants who received the intervention would not have developed NEC even if they had not received the intervention. Thus, 90% of NICU infants would receive the intervention unnecessarily. Any risk associated with the intervention is thus experienced by many infants who are at negligible risk of NEC. With accurate prediction models, fewer low risk infants would need to be exposed during trials of new preventive strategies.

A critical gap in knowledge, therefore, is the role of clinical risk factors in the susceptibility of premature infants to necrotizing enterocolitis. However, the potential clustering of NEC cases may compound the difficulty in predicting NEC among premature infants. Some have suggested that there may be differences between cases that occur in a cluster versus cases occurring randomly or sporadically, however very few have investigated this hypothesis.
Thus the type of case, clustered or sporadic, should ideally be considered in modeling risk factors.

The long term goal of this research is to identify a set of risk factors that define infants who are at significantly increased risk for necrotizing enterocolitis with the eventual goal of tailoring intervention strategies to prevent the onset of necrotizing enterocolitis in these high risk infants. The current study is the first stage in achieving this long-term goal, through the characterization of associations between physiologic traits and necrotizing enterocolitis in a large sample of very low birth weight infants admitted to the neonatal intensive care unit. The overall objective of this study is to explore the relationships between medical and physiologic characteristics and NEC disease in infants, and better understand how these relationships may be used to clinically predict the risk of NEC in this population.

**Hypotheses and Specific Aims**

The central hypothesis of this proposal is that a set of identifiable risk factors for NEC exist among VLBW infants in the NICU, that clustering of cases occurs, and that the risk factors for NEC differ between cases that occur in a cluster and those that occur sporadically. In order to address the identified gaps in the knowledge base, the current study has four specific aims and several related hypotheses:

**Specific Aim 1:** Create and validate a model for predicting necrotizing enterocolitis in a population of very low birth weight infants utilizing physiologic, clinical, and biochemical parameters and compare the predictive value of the model to individual risk factors.

**Hypothesis 1:** Prediction of necrotizing enterocolitis will be significantly improved by modeling clinical, physiologic, and biochemical parameters when compared to using each factor alone to identify infants at risk for developing the disease.

**Hypothesis 2:** The multivariable associations of the predictors with the outcome in an external sample would not differ significantly from those in the original internal sample.
**Specific Aim 2:** Create a regression or classification tree using variables from the prediction model and compare the two statistical methods with regards to the sensitivity of each model’s ability to identify patients at risk for NEC.

**Hypothesis:** Binary recursive partitioning will be more sensitive than logistic regression in correctly classifying patients at risk for necrotizing enterocolitis.

**Specific Aim 3:** Examine the clustering of NEC cases within a NICU setting.

**Hypothesis:** Necrotizing enterocolitis occurs as sporadic, unrelated cases and as groups of cases that cluster over time.

**Specific Aim 4:** Examine risk factor differences between sporadic cases and clustered cases of NEC.

**Hypothesis:** The risk factors for developing the disease will differ between cases that are considered unrelated or sporadic and cases that are identified as occurring in clusters.
Chapter 2: Background

Epidemiology of Necrotizing Enterocolitis

Necrotizing enterocolitis (NEC), the most common gastrointestinal emergency in the neonatal intensive care unit (NICU), is a sudden-onset life-threatening disease that occurs primarily in premature infants. This disease of neonates was first described as case reports of gastrointestinal perforation in the late 19th century[1]. In 1939, Thelander and colleagues assembled a review on 85 cases of unexplained ruptured viscus in infants <1 year of age[2]. However, it was not until 1953 that the name necrotizing enterocolitis was applied to this disease[3, 4] and by the 1960s Berdon[5] and Mizrahi [6] and respective colleagues had characterized it in depth, including clinical and radiological details. Since the1970’s, NEC remains the leading indication for emergent gastrointestinal surgery in neonates.

NEC is characterized by gastrointestinal and systemic signs and symptoms including feeding intolerance, delayed gastric emptying, abdominal distention or tenderness, occult or gross blood in the stool, lethargy, apnea, respiratory distress, and poor perfusion [1, 7]. Although the diagnosis is suspected from clinical presentation, it must be confirmed through diagnostic radiographs, surgery, or autopsy. In 1978, Bell et al proposed a system for categorizing NEC by presentation and severity [8]. This clinical staging system was later modified by Walsh and Kliegman [9] to include systemic, intestinal, and radiographic signs and to suggest treatment based on stage and severity of illness (Table 1).

The incidence of NEC is reported to be 2 to 5% among premature infants and 7-13% among VLBW infants[10-13]. The case-fatality rate is as high as 20-40% [14]. Few population-based studies report the number of cases in an entire state or geographic region over time, as most descriptive studies of NEC report numbers of cases in a single institution for a designated time period. Population-based studies report the incidence rate of NEC to be between 0.3 to 2.4 cases per 1000 live births[15-18]. Stoll extrapolated this data to all live births in the United
States each year, stating that with approximately 4 million births each year, 1200 to 9600 newborns are estimated to develop NEC in the United States each year [1].

The age of onset of NEC in the premature infant tends to be between 1 and 3 weeks of life, rarely occurring in the 1st 3 days of life or beyond 3 weeks of life[19]. The disease process often begins in the distal small intestine and presents suddenly with abdominal distension, signs of infection, unstable vital signs and cardio-respiratory collapse [20], yet the cause and/or causes of NEC are unknown. Neonates with NEC present with bloody stools, feeding intolerance, and bowel mucosa damage due to intestinal ischemia [13].

**Complications and Death due to Necrotizing Enterocolitis**

NEC in preterm infants is associated with a significant risk of death or long-term morbidities such as short-gut syndrome and associated cholestatic liver disease. Approximately 70% of patients require surgery, with complications contributing to morbidity and mortality [1]. Post-NEC strictures found in the colon may occur in 39%. Recent studies have shown an increase in the rate of poor neurodevelopmental outcomes for infants who survive NEC. Sonntag et al found a significantly higher rate of mental retardation among 20 infants who had NEC compared with age matched controls [21]. In a case-control study conducted by Sahlab et al, infants weighing < 1000 g at birth who had NEC were more likely than controls to demonstrate neurodevelopmental delay at 18 months of age, specifically affecting psychomotor function (OR: 18.5 95% CI: 1.9, 182.3) [22]. This is consistent with findings reported by Yeh and colleagues, who found that at 18 months of gestation-corrected age, infants who had NEC were more likely to have delayed psychomotor development compared to birth weight matched controls (OR:4.4 95% CI: 1.2, 17.1) [23]. Hintz et al found NEC to be a significant predictor of neurodevelopmental morbidity among infants enrolled in the NICHD Neonatal Research Network [24]. This finding was independent of birth weight and devastating complications such as intraventricular or periventricular hemorrhages.
Pathogenesis and Risk Factors

Necrotizing enterocolitis is considered a multifactorial disease with three major factors playing a role in the pathophysiology: enteral feeding, vascular or perfusion related gastrointestinal compromise, and bacterial invasion [25, 26]. These factors may somehow come together to produce bowel injury [26]. Figure 1, adapted from Noerr 2003[27], summarizes the factors that have been reported to play a likely role in the development of NEC[13, 19].

With the advancements in neonatal medicine, such as mechanical ventilation and surfactant therapy, more extremely premature infants are surviving well beyond the neonatal period, increasing the number of infants at risk for NEC[1, 19, 28]. While many studies have identified individual risk factors related to the development of NEC, most studies include only a small number of neonates with NEC, are single-institution reports, or were conducted in the presurfactant era [29-36]. Authors focused on a single factor, rather than exploring the additive or multiplicative effects of several factors. Table 2 is a summary of the studies that have focused specifically on identifying factors associated with an increased risk of necrotizing enterocolitis among premature infants. Most studies cited infants with birth weights ≤ 2000 g [37-41], while others cited gestational ages < 38 weeks[42-45]. A common consensus is that general prematurity is the most common risk factor; extremely low birth weight infants having greater severity of disease and mortality [42, 46]. Bacterial colonization has been implicated in the pathogenesis [47], though not one specific pathogen has been consistently identified, nor has a pathogen always been linked to cases of NEC[48]. It has been stated that frequency of use of antibiotics among the preterm population can alter the normal flora, hence leading to colonization of resistant strains of bacteria [49].

At least 80% of NEC cases are preterm/very low birth weight infants, with the incidence of the disease inversely proportional to the gestational age[42, 50]. Uauy et al reported an increased incidence of NEC among VLBW infants associated with prolonged rupture of
membranes, maternal age, birth weight < 1000 g, 5-minute Apgar score < 7, and maternal
(antepartum) hemorrhage, while finding a protective association with regards to high maternal
blood pressure[40]. Results reported by Uauy and colleagues also suggested a potential
interaction between gender and race. Guthrie et al showed that decreasing birth weight, low 5-
minute Apgar score, and mechanical ventilation requirements to be important risk factors in a
logistic regression model[42]. In a study reported by Guillet and colleagues focusing on H2-
blocker use and NEC risk, results from unadjusted analysis of the whole population found an
association between increased odds of NEC among African American infants, infants with
decreasing birth weight, and infants who were outborn[41].

Patent ductus arteriosus (PDA) has been shown to be associated with an increased risk
of NEC since PDA can cause diminished perfusion of the gastrointestinal tract [51, 52] [53].
Patients treated with indomethacin to close PDA, a common problem among premature infants,
may have an increased risk for NEC, possibly due to the effect of the drug on intestinal
circulation [13, 53]. This relationship, however, has not been consistently found in all studies.
Early studies (pre-surfactant era prior to mid-1990's) specifically designed to look at risk factors
for NEC failed to find an association between NEC and PDA, regardless of treatment [44, 45],
while later studies reported an increased risk among infants who received treatment for PDA[38,
43]. Advances in the supportive care of premature babies, (e.g. surfactant use, improved
technologies for mechanical ventilation) enable the very low birth weight infants to survive, and
thereby increase the population of patients susceptible to NEC.

**Birth weight and NEC**

Although there is a sharp decrease in the incidence of NEC in infants with a
postconceptual age greater than 36 weeks[54], and an inverse relationship is seen between
NEC and gestational age or birth weight[40, 45, 55], few studies have had the opportunity to
investigate differences in susceptibility for NEC based on birth weight categories. Investigators
have suggested that necrotizing enterocolitis may be a different disease in full-term infants compared to preterm infants, and in many studies of NEC, prematurity seems to be the main consistent risk factor identified. Beeby et al conducted a matched case-control study including newborn infants between 1984 and 1991 (pre-surfactant era) at the University of Sydney[45]. They reported that premature infants less than 29 weeks gestation had no more risk factors than their gestation-matched controls. Infants between 30 and 36 weeks with NEC, however, were more likely to have been small for gestational age or had evidence of asphyxia (based on low Apgar scores). It wasn’t until later that Luig and colleagues reported that different patterns of susceptibility based on gestational age grouping remained even during the surfactant era. Results reported by Luig indicated that, apart from surgical PDA, there were no significant factors associated with greater risk of NEC among infants less than 28 weeks gestation [43]. Infants who were between 28 and 31 weeks gestation had a number of perinatal risk factors associated with increased NEC risk, including lower birth weight, younger maternal age, placental abruption, hypertensive disease of pregnancy, and PDA.

Feeding and NEC

In 1975, Santulli et al suggested that enteral feedings played a role in the pathogenesis of NEC [25]. To date, studies implicating an etiologic relationship between feeding and NEC are controversial; NEC is the dominant argument for postponing enteral feeding, though it occurs in infants who are fed parenterally (receiving no oral feeds). The association between enteral feeding and NEC has been historically investigated in four main ways: the amount of feeds, the timing of initiation of feeds, the rate of advancement of feeds, and the type of milk fed. All studies that have looked at the role of feeding are highly variable with definitions, amounts, and timing of feeds. Tyson and Kennedy have performed a series of systematic reviews, published in the Cochrane Database of Systematic Reviews, looking specifically at three of the main
topics around the relationship between enteral feeding and NEC: minimal enteral, early versus delayed, and the rate of advancement[33, 56, 57].

There is no general agreement about the optimal timing to start early enteral feeds. Because enteral feeding promotes gut maturation[58-61] and benefits the immune function[62], there are potential clinical benefits. Early enteral feeding is suggested to enhance priming of the gut making it more apt to absorb nutrients, [63, 64] thus providing overall nutritional benefit to these vulnerable infants. A meta-analysis of 9 studies concluded that there was no significant effect on necrotizing enterocolitis comparing minimal enteral feeds with no feeds (RR 1.16 [95% CI 0.75, 1.79]).

Initiation of enteral feeds early without a delay in advancing feedings is an alternative feeding strategy. The evidence supporting the benefits and risk of early versus delayed initiation feeding strategies is insufficient. A meta-analysis conducted in 2004 reported only one study which focused on early feeding strategies and the risk of NEC, finding no significant relationship between the two (RR 0.53 [95% CI 0.11, 2.70]) [56]. LeGamma et al studied 20 infants with delayed feedings (no feedings for 2 weeks) and compared them to 18 infants who received feeds during the first 2 weeks of life [65]. Investigators described a higher incidence of NEC in the delayed group (60%) compared to the early-oral-feeding group (22%), and suggested that withholding oral feeds did not lower the incidence of NEC.

Limited data exist from randomized trials of feeding preterm infants comparing formula with human milk. Lucas and Cole reported a higher incidence of NEC among formula fed infants compared to human milk fed infants (OR 2.5 [95% CI 1.2, 5.2]) [66]. Schanler et al reported fewer cases of NEC among infants fed fortified human milk compared to preterm formula (1.6% vs. 13%, p< 0.01) [67]. Findings from a meta-analysis on donor human milk versus formula found that infants who received donor human milk were four times less likely to develop confirmed NEC compared to infants who received formula (RR 0.25 [95% CI 0.6, 0.98]) [68]. Schanler et al later reported on the effects of donor milk versus preterm formula in a
randomized trial with 243 extremely premature infants and found slightly decreased rates of NEC among infants fed donor milk compared to infants fed preterm formula, though this finding was not statistically significant (6% vs. 11%; p=0.27) [69]. However, the investigators concluded that beneficial short-term outcomes for extremely premature infants are not supported by the substitution of pasteurized donor milk for mothers own milk. Unpublished data by Meinzen-Derr et al suggests that the amount of human milk fed in the first 2 weeks may play an important protective role against NEC or death among infants who are extremely low birth weight (< 1000 g) [70].

It appears that NEC among very low birth weight infants is the result of a multi-faceted etiology, with prematurity the primary and consistent factor most frequently associated with the disease. There is a great need for a better understanding of this complex disease, including identifying those factors that might put an infant at the highest risk, so that individualized clinical care may be improved and potential intervention strategies may be targeted to those with greatest need.

**Epidemics and outbreaks**

According to the National Institute of Communicable Diseases, an outbreak or epidemic is defined as the “occurrence in a community of cases of an illness clearly in excess of expected numbers “[71]. While an outbreak is usually limited to a small focal area, an epidemic covers larger geographic areas and has more than one focal point. The number of cases which are needed to be called an outbreak is dependent upon several factors. Historical patterns of the disease, case fatality, and potential of spread to other areas (the infectiousness of the disease) play a role in defining an outbreak for a particular disease. For some diseases (i.e. poliomyelitis), a single case would comprise an outbreak.

In 1965, Mizhari et al noted the lack of NEC epidemics and considered this as evidence for nontransmissability [6]. However, in 1972, 2 outbreaks of NEC among neonates were
reported in South Africa and again in India in 1973. The first outbreaks of NEC in the United States were reported by Virnig et al in 1974 and then soon after in 1975 by Desai et al [72, 73]. Although NEC generally occurs as sporadic cases, “epidemics,” or outbreaks, continue to be reported[48].

Outbreaks of NEC have been reported in association with a variety of infectious agents, included *Escherichia coli*, *Klebsiella*, *Enterobacter*, *Pseudomonas*, *Salmonella*, *Clostridium perfringens*, *Clostridium difficile*, *Clostridium butyricum*, coagulase-negative Staphylococci, coronavirus, rotavirus, and enteroviruses [3, 74-81]. Although a variety of studies suggest an infectious pathogen is a potential cause for the observed clustering of this devastating disease, failure to identify a single etiologic agent makes it difficult to determine cause and appropriate prevention strategies. Gupta, et al, for example, conducted a comprehensive analysis of bacterial, parasitic and viral agents present in stool samples of 23 NEC cases and 33 controls in a NICU and found no evidence that endemic NEC within their institution was caused by a single agent, nor was there a suggestion of multiple small outbreaks having one or more infectious agents in common [82].

Regardless of the paucity of consistent associations with specific pathogens, NEC “epidemics” exist in the literature. Table 3 describes the characteristics of 21 outbreaks or clusters of NEC reported in the literature. It is unclear if the Bell staging case definition was applied in all investigations, and some investigations included both suspected NEC and gastroenteritis cases. Most of the epidemics reported a duration of 8 to 10 weeks[79, 83-93], though one reported a duration of 20 weeks[85] and one for as long as 10 months[94]. Most of the affected newborns had a low birth weight, though several of the studies reported NEC outbreaks in infants > 2000 g[85, 91, 95, 96]. Many of the investigations reported no apparent risk factors with regards to epidemic cases (Table 4). Many of the NEC cluster investigations reported risk factors associated with NEC, the most common of which were low birth weight, earlier gestational age[74, 81, 87, 89], low Apgar score[97], and perinatal complications[79].
Book, et al reported evidence that both temporal and geographic clustering of 74 NEC cases occurred within a newborn intensive care facility during a 42 month period; 69 of these cases occurred during a span of 10 months[94]. In 1979, Guinan and colleagues reported 3 “epidemics” in 3 different high-risk nurseries, all occurring within the same period[85]. Chaney et al reported on temporal clustering of NEC occurring in 2 different maternity hospitals in Paris in 1982[86], while Faustini et al, in 2004, reported simultaneous temporal clustering of NEC within 2 different neonatal units [96].

Clusters of Necrotizing Enterocolitis

The term “cluster” may be defined as the occurrence of a group of cases in a circumscribed place and time, in amounts that are considered (by the public or others) to be greater than expected. The cluster is usually based on anecdotal evidence, and often the first task is to determine whether the number of cases truly is or is not greater than expected[71]. There have been many anecdotal episodes of temporal clusters of NEC [48, 96], but no infective agent consistently has been linked to its occurrence [82]. Although NEC may show nosocomial spread [48, 95, 98], and fits part of the Centers for Disease Control and Prevention’s definition for nosocomial infections [99], no formal demonstration of an infectious process has ever been made. To date, few, if any, studies have adequately addressed the identification of NEC clusters nor proposed formal statistically reliable methods for detecting NEC clusters.

Given the severity of NEC, it is fundamental that healthcare workers in neonatology and neonatal intensive care units are made aware of the potential for clusters to occur and that all necessary preventive measures be taken. Differentiation between clustered and sporadic NEC may enable clinicians to identify risk factors obscured when these cases are combined for analysis.
**Prediction Models**

Clinical prediction rules are tools designed to assist medical decision making and are intended for use by clinicians when caring for patients. Specifically, predictive modeling uses statistical models to estimate the likelihood of an outcome for a given patient from a set of observations that are particular to that patient [100, 101]. As these rules or models are used to make decisions about patient care, it is important that they be well developed and validated.

When empirical data are available, individualized probability estimates for such outcomes may well be obtained from a logistic regression model, especially if such a model is developed in a large data set [102]. Advantages of logistic regression include: 1) the parameter of interest can be bounded between 0 and 1 (binomial); and 2) interpretable results are provided, because the regression coefficients represent log-odds ratios [103]. By conducting multivariable analyses through logistic regression, the relationship among a series of predictors with the outcome can be modeled, allowing for better understanding of these relationships. Regression modeling also allows for the testing of statistical interactions among independent variables, which assess differences in the effects of one or more independent variable according to levels of another independent variable, though the interpretation may become difficult when many variables are assessed.

Predictive modeling through multivariable logistic regression analyses appear throughout the pulmonary and critical care literature [104-108]. To avoid performing unnecessary diagnostic tests or to help improve the clinical management, several prediction/decision rules have been developed specifically for the pediatric population. Oostenbrink and colleagues developed a decision rule to help determine whether to start empiric antibiotics in children with signs of meningeal irritation[109] while simultaneously developing a model predictive for the occurrence of permanent neurological sequelae or death after bacterial meningitis[110, 111]. Nigrovic and colleagues found a multivariable model to be accurate in identifying children at high risk for bacterial verses viral meningitis[112]. Lynch and
colleagues developed a model used to help identify those children at risk for pneumonia seen in the emergency department[113] in order to help guide clinicians with regards to necessary diagnostic chest radiographs. Other models, as well, have been developed with the idea that having a set of predictors will help eliminate performing unnecessary and potentially invasive diagnostic tests [114, 115]. Other examples of the clinical utility of predictive modeling are seen by the use of the newborn illness severity and mortality risk scores, which were created and validated through predictive modeling techniques [116-119].

In neonatology, one example of a clinical prediction tool is that developed for determining, in preterm infants, risk of retinopathy of prematurity (ROP).[120] Although developed for clinical purposes, this ROP model was subsequently used to enroll higher risk infants in a trial of early retinal ablation[121]. To date, a model developed for the purpose of predicting NEC among infants in the NICU does not exist.

**Importance of Model Validation**

Logistic regression may be used to develop predictive models for dichotomous outcomes, such as short-term mortality. When a previously developed model is applied in another center and/or in a more recent time period, the external validity (or generalizability) of model predictions is addressed [122-124]. Successful validation is not completed until the model is validated on an independent external data set (the external validation data set). External validation aims to address the accuracy of a model in patients from a different but plausibly related population. Assessing the external validation of a model allows us to consider updating a previously developed model, such that the predictive model is adjusted to local and/or contemporary circumstances [122, 123, 125-127].

**Recursive Partitioning**

Recursive partitioning, also known as classification and regression tree analysis (CART), is a nonparametric statistical procedure that identifies mutually exclusive and exhaustive
subgroups of a population whose members share common characteristics that influence the
dependent variable of interest [128]. A classification tree can be viewed as a decision rule for
assigning items to groups using a sequence of binary splits for the explanatory variables[129].
Tree-based models arising from the CART methodology provide an alternative to linear and
additive models for regression problems and to linear logistic and additive logistic models for the
classification problem. CART analysis uses recursive partitioning and asymmetric stratification
to develop tree-based models using binary predictors [129-131]. It uses splitting rules to stratify
data into risk groups, and then considers further splits among the subgroups already included in
the model. CART analysis can be used simply to explore the data, identify possible high-risk
subgroups, and uncover interactions or effect modifications among prognostic factors.

A generic illustration of CART output is presented in Figure 2. Classification and
regression trees begin with one “node”, or group, containing the entire sample, called a parent
node. The CART procedure examines all possible independent, or splitting, variables and
selects the one that results in binary groups (child nodes) that are most different with respect to
the dependent variable, according to a predetermined splitting criterion. The ideal split would
divide a group into two child groups in such a way so that all of the rows in the left child node
have the same value on the target variable and all of the rows in the right group have the same
target value, but different than the left group[128]. However, because such a perfect split is
rare, the general aim of is to achieve the best possible predictive accuracy, defined as the
prediction with lowest misclassification rate. The algorithm commonly used for CART analysis
is the Gini index [132]. This index reaches a value of zero when only one class is present at a
node. Each split is chosen to maximize the heterogeneity of the categories of the target
variable in the child nodes.

As applied to case-control data, each binary split in a classification tree yields two
subgroups, one that contains a relatively high proportion of cases and the other that contains a
relatively high proportion of controls. The combination of these binary splits can be interpreted
as providing a set of prediction rules that may be used for classifying subjects according to the probability of being a case. Recursive partitioning has been applied in several prognostic applications including the prediction of growth failure in infants [133], awakening from coma [134, 135], fatal outcomes in meningococcal disease in children [136], and preterm delivery and small for gestational age birth [137].

Although logistic regression provides explicit measures for statistical inference and measures of the strength of the association between each variable and the outcome, estimating the probability of NEC for each patient involves manipulating the logistic regression formula. This could be done at the bedside with a device (e.g. laptop) which would compute the equation and output a probability of disease. A tool that is easily understood, easy to administer, and is sensitive in its ability to assess the risk of NEC in very low birth weight infants is important not only from a clinical standpoint, but also has utility in the realm of research and clinical trials. CART may provide such a tool, as obtaining the results of an infant’s risk of disease does not require computing an equation once the analysis is complete. Little is documented about how sensitive CART is in classifying false positives, which could introduce bias. It appears for low incidence outcomes, CART may be prone to over classify non-cases as cases (false positives). It is important to keep in mind that logistic regression can be used to eliminate variables with little predictive ability for the outcome; CART analysis will attempt to utilize all of the variables available. The steps for both logistic regression and CART analysis are described in detail in the methods section.
Chapter 3: Methods

Description of Study Population

In 1986, the National Institute of Child Health and Human Development (NICHD), of the National Institutes of Health, initiated the Neonatal Research Network (NRN). This Network was established to conduct multi-center clinical trials and observational studies in neonatal medicine in order to reduce infant morbidity and mortality and promote healthy outcomes. The Generic Data Base Study (GDBS), began in 1987 as a registry of baseline and outcome data for very low birth weight infants (VLBW) (weighing 401-1500 g at birth), based on data collected in a uniform manner from neonatal intensive care units (NICUs) at 16 participating institutions across the United States [138].

This research required an analysis of retrospective data collected through the NICHD NRN, which includes all VLBW admissions to NICUs of University Hospital, Cincinnati Children’s Hospital Medical Center (April 1, 1991 to May 31, 2005) and Good Samaritan Hospital (September 1, 1995 to May 31, 2005). Trained personnel collect maternal and delivery data soon after birth and infant data for 120 days (with end points of discharge or death). Clinical information are entered into the registry using an electronic case report form incorporating real-time validity checks. Informed consent is not required for registry entry. To preserve patient confidentiality direct patient identifiers (e.g. patient name) are not collected.

Power Calculation

Power calculations were conducted utilizing a fixed sample size of 3816 infants available for analyses (described in Selection of Study Cohort below). In order to address Specific Aim 1, the power calculation was based on a comparison of the area under the ROC curves (AUC) of 2 predictive models (a model with the single best predictor vs. a model with all significant risk factors) [139]. For the purpose of this power calculation, the following assumptions were made: 1) the overall rate of NEC in this population is approximately 7% (preliminary results); 2) the AUC of the single predictor model will
be .65 and the AUC of the full risk factor model will be .80; and 3) our available sample size for the creation of the predictive model is 3435 [90% of the overall included data]. With an $\alpha = 0.05$ and available sample size, we will have over 90% power ($1 - \beta$) to detect a difference in AUC of 0.15 between the 2 prediction models.

Little is known about power calculations for recursive partitioning or classification and regression tree analysis (CART), which would address Specific Aim 2. Because it is often used as a “data mining” strategy, not hypothesis testing, sample size/power calculations are often not considered for recursive partitioning and the available software does not include any information or statistical methodology to tackle the issue of power. Neither Breiman or Zhang, two of the major contributors to what is known about CART analysis, discuss power or sample size needs for this type of analysis [130, 132].

Power calculations for Specific Aim 3 (clustering of NEC) were based on differences in incidence density rates between an expected rate of NEC and an observed rate of NEC. Assuming that the background, or expected, rate of NEC is 1.29 per 1000 patient-days, and using a 1-sided 1-sample Poisson distribution to test the null hypothesis that the observed rate of NEC during an identified cluster period is not significantly higher than the expected rate of NEC, we should be able to identify clusters when the incidence density rate during a specified period is at least 6.0 per 1000 patient-days. A 1-sided test was used for the power calculation for this hypothesis since we were interested in incidence rates that were specifically higher than the expected rate.

**Selection of Study Cohort**

All VLBW admissions (401 to 1500 grams birth weight) to the 3 NICUs listed above are registered with the NRN and are potentially eligible for inclusion. Infants were excluded based on the following criteria: (1) death < 24 hours of life; or (2) congenital anomalies of heart and gut malformations. Currently 4176 infants have been enrolled into the GDBS; 3906 met our
inclusion criteria and were available for analysis. Because 98% of this population is non-Hispanic White or African-American, there was not enough information to make inferences with regards to other races/ethnic groups. Therefore, the final analyses were limited to 3816 infants who were non-Hispanic White and African-American.

**Data Collection**

Data is routinely collected by research nurses/coordinators on all new admissions into one of three NICUs. Data includes clinical data (i.e. birth weight, sex, race, Apgar scores, gestational age, respiratory status, steroid and surfactant use), nutrition data (i.e. parenteral and enteral intake, days to regain birth weight, episodes of proven NEC), and maternal information (i.e., maternal age and race, marital status, parity, gravida, method of delivery). The date of birth was used to help establish the age of the infant at certain time points and an accurate length of stay in the NICU. Variables that may indicate the health status of the child were also evaluated (Apgar scores, respiratory status, intracranial hemorrhage, number of days on mechanical ventilation, number of days on parenteral nutrition, number of days to regain birth weight, number of days to reach full enteral feeds, feeding intolerance, growth data), and control for antibiotic prophylaxis, drugs, steroids, and surfactants that may have been administered.

**Case Definition**

Because studies have used a variety of definitions in order to define cases of NEC, this research used the definition of NEC as proposed by Bell’s staging for necrotizing enterocolitis[8]. (See Table 1 for the definitions of NEC.) For the purpose of this research study, NEC was defined as either Bell stage II (distension and/or signs of peritonitis with x-ray evidence of pneumatosis intestinalis) or III (NEC confirmed at surgery). The date of the first episode of NEC was also recorded, which was used to define an occurrence of an epidemic or cluster (Specific Aim 3). Although Bell’s stage I (suspect NEC cases) is of interest, beginning in
1998, infants with suspected NEC were no longer differentiated from infants who did not acquire NEC. Therefore, this study focused on NEC as defined as definite or confirmed (stage II and III).

**Overview and Justification of Statistical Methods**

*Dichotomized and Calculated Variables*

The only continuous variables that were considered for this analyses were birth weight (measured in grams), gestational age (measured in weeks), and maternal age (measured in years). Birth weight was also evaluated as a dichotomous variable (<1000 grams and > 1000 grams). This cutoff for birth weight was selected by identifying values associated with a point of the ROC curve with the best sensitivity-specificity trade off (Figure 3). This optimal cutoff point typically occurs where the ROC curve turns the corner, at which point each incremental gain in sensitivity results in a substantial loss of specificity [112].

Prolonged rupture of membranes (PROM) was defined as first evidence of rupture of amniotic membranes occurring more than 24 hours prior to delivery. PROM has been associated with increased risk of infection and the use of antibiotics, which may influence the gastrointestinal microbial colonization, a possible risk factor for NEC[140]. Patent ductus arteriosus (PDA) was defined as either clinical evidence of a left to right shunt (the presence of continuous murmur, hyperactive precordium, wide pulse pressure, bounding pulses, congestive cardiac failure, increasing oxygen requirement or evidence of increased pulmonary vascularity on x-ray) and/or echocardiographic evidence of a ductus with left to right shunt. Mechanical ventilation, a variable that may be a surrogate for illness severity of the infant, was defined as receiving any mechanical ventilation for more than 2 days within the first 10 days of life. Antepartum hemorrhage was defined as placenta previa, abruption or threatened abortion resulting in bleeding, documented after 20 weeks of pregnancy. Although there is insufficient evidence regarding the relationship between the amount of enteral feeding and NEC [67, 141,
142], or timing of the initiation of enteral feeding [33, 65, 143, 144], enteral feeding was investigated as a potential covariate/risk factor for NEC. Enteral feeding was defined in 2 ways: a) receiving the first enteral feed (initiation of feeds) within the first 3 days of life; and b) achieving full enteral feeds (first day supplemental IV fluids were less than 20 cc/kg/day) within the first 5 days of life.

Variable Diagnostics

Each continuous variable was examined for distribution patterns using SAS PROC UNIVARIATE. Due to the high correlation between birth weight and gestational age, and that both variables are considered highly predictive of the outcome, it was decided that birth weight, not gestational age, would be used in the final analyses. Very low birth weight cohorts are common for studying at-risk premature infants. By definition they typically have 100% small-for-gestational age (SGA) infants beyond 32 weeks gestational age [145]. These analyses are based on a registry of infants ≤ 1500 grams birth weight, many of whom are small for gestational age (<10th percentile based on Alexander fetal growth curves [146]) (Figure 4). A variable indicating those infants who were small for gestational age, based on the Alexander fetal growth curves, was evaluated as a potential confounder.

Several variables for PDA were evaluated: any diagnosis of PDA, any treated PDA (with either indomethacin treatment or surgery), and surgical PDA. Because these 3 PDA variables are highly correlated in nature, PDA was categorized into mutually exclusive groups: surgical PDA, PDA medically treated with indomethacin, and untreated PDA.

Feeding information was evaluated at the age of receiving the first enteral feeding and the time it took to achieve full enteral feeds. The purpose of creating a logistic model for NEC is to be able to assist clinicians in predicting those infants who are likely to develop or acquire the disease, and to do so during the infant’s early days of their NICU stay. Estimating the probability of disease and treatment-associated harm could aid in the selection of individuals
who would benefit most from participation in trials of preventive interventions [147]. Therefore, variables for feeding (age at initiation and age at achieving full feeds) were dichotomized at times that made statistical sense (based on means or percentiles) and that made sense from an early clinical standpoint (time prior to introduction of potential intervention/clinical trial enrollment).

Unadjusted analyses between NEC outcome and categorical variables were completed using contingency table analysis and odds ratios to determine differences in distributions of maternal and infant clinical characteristics by the outcome (NEC vs. No NEC). This bivariable analyses was conducted using SAS PROC FREQ with the CMH option for obtaining odds ratios with 95% confidence intervals. The Fisher’s exact test was used when individual cells had fewer than five expected occurrences. Continuous variables were reported using mean and standard deviation or median and range for skewed distributions. Differences between infants with NEC and those without were tested using Student’s t-test (SAS PROC TTEST) or Wilcoxon Rank Sum test (SAS PROC NPAR1WAY with WILCOXON option) for continuous variables as appropriate.

**Logistic Regression Modeling**

Logistic regression modeling using PROC LOGISTIC was used to determine the contribution of certain clinical characteristics on the odds of NEC. Dependent variables were coded as 1 when the characteristic was present and 0 when the characteristic was absent. The DESCENDING option was used to model the risk of NEC. Due to the large number of potential risk factors for NEC, logistic models were developed using a stepwise variable selection procedure, with the criterion for entry into the model at p=0.20 and the criterion for staying in the model at p=0.15. For birth weight, odds ratios were calculated based on 100 gram unit change in birth weight. Once the model was fitted utilizing the stepwise selection, the model was refitted with the same variables using a backward selection method performed manually (not
computer selection). The variables were entered into the model based on the statistical significance of the Wald statistic. Interaction terms were investigated in logistic models, particularly with the birth weight and race, since both may have modified the effect of potential risk factors for the outcome. Further, stratified models were developed by birth weight (< 1000 grams and ≥ 1000 grams) to evaluate if different predictors of NEC exist by birth weight category; infants who are < 1000 (extremely low birth weight) are considered to be at higher risk for NEC [1, 40, 45, 55].

Model Diagnostics and Goodness-of-Fit Statistics
The LOGISTIC procedure produces a variety of statistics that can help with the evaluation of models. The Hosmer-Lemeshow (HL) test statistic was used to assess the goodness of fit of the models[148], implemented with the LACKFIT option in the MODEL statement. The predicted probabilities were generated based on the estimated model sorted by size, and grouped into 10 intervals. The expected frequency within each interval was then obtained by adding up the predicted probabilities, and then compared with the observed frequencies by the Pearson chi-square statistic, with degrees of freedom (DF) equaling the number of intervals minus 2 (8 DF for 10 intervals). A high p-value indicates a model with good fit. However, it is important to note that the HL goodness-of-fit test is considered a fairly conservative test and is highly dependent on how the observations are grouped. The Deviance statistic, along with examination of residuals and deviances, was used to evaluate the potential for overdispersion of the data. The AGGREGATE and SCALE=NONE options were used to get a deviance that has a chi-square distribution, without adjusting the goodness-of-fit statistics for overdispersion[149]. The AGGREGATE option calculates the Deviance and Pearson’s chi-square, which tests the null hypothesis that a better model does not exist. A significant chi-square leads to the rejection of the null and subsequent rejection of the current model. In other words, similar to the HL goodness-of-fit test, a high p-value indicates a relatively good model fit.
Regression diagnostics are used to indicate observations that may have undue influence on the model fit, or which may be outliers. Overly influential observations may lead to increased variance of predicted values\[150\]. Several methods exist to measure influence. The Pearson (DIFCHISQ) and Deviance (DIFDEV) residuals identify observations that are not well explained by the model. The INFLUENCE option displays the values of the predictor variables for each observation, a column for each diagnostic produced, and the observation number, and the measures are graphed to highlight extreme values. Plots of the change in the deviance and the Pearson chi-square by predicted probabilities were evaluated, to determine which observations needed to be examined as potential outliers. According to Hosmer and Lemeshow, because it is a crude approximation of the 95th percentile of the chi-square distribution with 1 degree of freedom, values above 4 indicate observations with unusually large residuals, which may influence the parameter estimates [148]. Because no published statistical cutoffs exist for logistic regression diagnostic statistics; a cutoff of the 95th percentile for the distribution of DIFDEV, DIFCHISQ, H and C were established. If any observation had a diagnostic statistic greater than or equal to the 95th percentile, then the observation was deleted and the model was rerun. Observations with high DFBETAS were also examined in the same manner. DFBETAS is the change in the vector of regression coefficient estimates upon deletion of each observation in turn, scaled by their standard errors[150], and hence assesses the effect of an individual observation on the estimated parameter of the fitted model. If the deletion of a few observations substantially altered the values of the parameter estimates, then the value of that variable selected for the model was re-evaluated\[151\].

**Recursive Partitioning**

Recursive partitioning software was used (DTREG version 4.5, Phillip H. Sherrod) to create the decision tree. Variables that were significantly associated with NEC (p<0.20) were entered into the model, which included the same candidate variables for the logistic model. The
Gini method for classification trees and 10-fold cross-validation was used. Recursive partitioning methods start with the entire population, which is represented by the first node and a sequence of binary separations (splits) of a group of subjects; at each split, all included predictive variables are examined. The subgroups that result from the splits are called child nodes. A classification tree is formed by sequentially splitting the nodes of previous splits until some specified stopping criteria are met; the final nodes are called terminal nodes (Figure 2). The numerical objective of partitioning is to make the contents of the terminal nodes as homogenous as possible. A quantitative measure of the extent of node homogeneity is the notion of node impurity. The homogeneity of a node is based on the prevalence rate of an event within that node. An example of node impurity can be seen as:

\[
\text{Number of infants with NEC} \\
\text{Total number of infants in the node}
\]

The closer this ratio is to 0 or 1, the more homogeneous is the node [132]. A node with no impurity would have no variability in the dependent variable within that node (all 0s or all 1s). The highest amount of impurity is when \( p_{ij} = 0.5 \), where \( p_{ij} \) is the probability that the dependent variable (in this case NEC) is equal to \( i \) in Node \( j \), where \( i \) can take values 0 or 1 [128]. The splitting criteria are based on functions of \( p_{ij} \), known as impurity functions, which select the split that has the largest difference between the impurity of the parent node and a weighted average of the impurity of the two child nodes.

Without specified stopping criteria for tree size, a tree might be created so large that every observation will have its own terminal node, making the tree so large that it would be difficult or impossible to interpret. For simplicity, nodes containing fewer than 10 observations were not split. DTREG uses what is known as backward pruning in order to create a tree of optimal size. This method is the suggested method for pruning trees[150]. Beginning at the last level, or the terminal nodes, the child nodes are pruned away if the resulting change in predicted
misclassification cost is less than the complexity measure. This measure indicates how much additional accuracy a split must add to the entire tree to warrant the additional complexity.

Various approaches have been proposed to deal with missing data. Zhang and Singer suggest replacing missing data with the minimum or maximum values and examine the corresponding entropy results[132]. Breiman and colleagues suggest using surrogate splits for subjects with missing data [130]. Surrogate splitters are predictor variables that are not considered as good at splitting a group as the primary splitter, but which yield similar splitting results; in essence they mimic the splits produced by the primary splitter. The surrogate splitter approach suggested by Breiman et al was adopted for the purpose of these analyses.

**Detailed Analysis Plan by Specific Aim**

**Specific Aim 1**

The purpose of Specific Aim 1 was to create and validate a model for predicting necrotizing enterocolitis in a population of very low birth weight infants utilizing physiologic clinical, and biochemical parameters. To achieve this aim, the database was randomly divided into a derivation sample comprising of 90% (n=3435) of the patients and a validation sample comprising of the remaining 10% (n=381). Because the proportion of infants with NEC was fairly low (approximately 7%), and since strong validation procedures were implemented (i.e. external validation described later in this chapter), the decision to split the data at 90%/10% was made. This split allowed us to maximize the model derivation procedures. It is understood that other splits (e.g. 50/50) might have been considered more optimal. The Student t test was used to compare means and the chi-square was used to compare differences in proportions between the derivation and validation samples and to identify variables in the derivation set that were marginally significantly associated (p<0.15) with NEC. Potential predictors were identified through what is known in the literature, what is known clinically, and those variables that not
only have been found to be associated statistically with the outcome during bivariable analysis, but also are biologically plausible. As described earlier under Variable Diagnostics, because of high correlation between birth weight and gestational age, birth weight was selected.

**Model Development-Logistic Regression Modeling**

The logistic regression model predicts the probability of an event (NEC), conditional on infant and disease characteristics. Using logistic regression models, the predicted risk of NEC was calculated for an individual with any combination of demographic or disease characteristics. In model selection, the maximum-likelihood test statistic is used to determine the statistical significance of the predictor variables.

The SAS PROC LOGISTIC procedure, CTABLE and PPROB output were used to calculate the sensitivity, specificity, and overall classification correctness for the model. Models with interaction terms between birth weight, race, and important predictors were constructed. The Akaike’s Information Criterion (AIC) was used to compare the relative fit of different models (i.e. models with and without interaction terms). The AIC is calculated as \(-2 \times \text{log-likelihood} + 2k\) where \(k\) is the number of estimated parameters[149]. As a general rule of thumb, lower values of the AIC correspond to more desirable models.

The variables that were evaluated for this analysis were listed earlier in this section. In summary, variables considered in the logistic regression analysis, with NEC as the dependent variable, included maternal factors such as race, age, parity/gravidity, marital status, prenatal care (defined as at least one prenatal visit), hypertension, diabetes mellitus, multiple gestation, mode of delivery, prolonged rupture of membranes, maternal hemorrhage, and antenatal steroids. Infant factors included gender, birth weight, small for gestational age, Apgar score at 5 minutes of less than 7, early antibiotic administration, and feeding information. As stated under Variable Diagnostics, gestational age was not included in the logistic regression because it is highly correlated with birth weight and populations of VLBW infants tend to over-represent
children who are small for gestational age. Once the models were created using the
STEPWISE selection procedure, they were recreated with backward selection (done manually,
not automatically by computer) to verify the findings. All variables were included at first; the
variables with the highest p-values (lowest chi-squares) were removed one by one and the
model fit was recomputed.

**Evaluation of Model Performance**

**Discrimination and Calibration**

The evaluation of model performance focuses on discrimination and calibration.

Discrimination refers to the ability to distinguish high-risk from low risk patients. The model’s
derivation was tested via the receiver operator characteristic (ROC) curve. The ROC curve
graphically represents the relationship between the sensitivity (probability of correctly classifying
a positive case) and the specificity (probability of correctly classifying a negative case) of the
model. This curve is useful in selecting specific thresholds on which predictions might be
based. A single-valued measure of test accuracy is the area of the graph lying beneath the
curve. A test with a ROC curve of 0.50 is considered to be achieved by chance alone[150]. A
test with a ROC curve of 1.0 is considered to have perfect discrimination.

Calibration refers to whether the predicted probabilities agree with the observed
probabilities. Several goodness-of-fit statistics are available to quantify calibration. For the
purpose of this proposal, calibration was assessed by calculating the Hosmer-Lemeshow
goodness of fit statistic and plotting the predicted probabilities against the observed
probabilities.

**Internal Validation**

After the development of the multivariate risk prediction model utilizing the derivation
sample (n=3435), the assessment of its goodness of fit and discrimination, the model was then
applied to the validation sample (n=381) in order to test the discrimination on an independent
sample. The multivariable coefficients for each risk factor in the derivation sample were compared to those of the validation sample by using t tests. In order to test the sub-hypothesis that the overall multivariable effects of the predictors did not differ between the derivation and validation models, the log-likelihood of the derivation prediction model was compared to the log-likelihood of the validation model with the same predictors but with re-estimated coefficients. Once the model was tested on the validation sample, the data from both samples were combined and the model was re-evaluated.

Comparison of the Final Model to a Simple Model

In order to fully address the hypothesis that a multivariate model is more predictive of NEC than a simple model (containing only one predictor), the area under the ROC curve for the final multivariate model was compared to simple models, each containing a single predictor, via methods reported by Hanley and McNeil [152]. The method proposed by Hanley and McNeil refines the statistical comparison of the areas under two ROC curves derived from the same set of subjects by taking into account the correlation between the areas that is induced by the paired nature of the data. The general approach to this comparison is defined as:

$$ z = \frac{A_1 - A_2}{\sqrt{(SE_1^2 + SE_2^2 - 2rSE_1SE_2)}} $$

where $A_1$ and $SE_1$ refer to the observed area and estimated standard error of the ROC areas associated with model 1; where $A_2$ and $SE_2$ refer to corresponding values for model 2; $r$ represents the estimated correlation between $A_1$ and $A_2$.

External Validation

To determine the generalizability of the derived prediction model, the model was applied to an external data set (n=32,770). This external sample included data collected on very low birth weight infants during the same time period (January 1991 – December 2004) from 16
NICUs participating in the Neonatal Research Network. Inclusion and exclusion criteria were identical to those applied to the internal sample. Data collection and definitions of diagnostic determinants were identical to the internal sample.

External validation of this model occurred in two ways: 1) the regression coefficients were fixed, and the logistic regression equation was applied to external sample to determine the sensitivities and specificities of the current prediction model on a new but similar population; and 2) the model with the same variables were applied to the external sample and the regression coefficients were re-estimated. The first method allowed for evaluation of the current prediction model on a new population and the second method allowed for evaluation of the variables chosen for the model. Once the prediction model was applied to infants in the external validation set, the performance (ROC area) of the model as well as the calibration was tested. Calibration curves were constructed by plotting the estimated NEC proportions stratified by 10% intervals (x-axis) against the observed NEC proportions (y-axis). Overall, the sub-hypothesis of this aim was that the multivariable associations of the predictors with the outcomes in the external validation sample would not differ from those in the internal derivation sample.

**Specific Aim 2**

The purpose of Specific Aim 2 is to create a classification tree using the candidate variables described for the prediction model created in Specific Aim 1 (i.e. the same dichotomous predictors used in the logistic regression analysis), and to compare the results of the classification tree to the results from the logistic regression analysis. This analysis begins with the entire study population, which is first divided into two subpopulations, each of which is as homogeneous as possible. The division of the population of VLBW infants was carried out through one covariate while the homogeneity is measured by the distribution of the outcome (NEC) in the subpopulation, and this process was repeated until the stopping rules were met.
The two statistical methods (logistic and recursive partitioning) were compared with regards to the sensitivity of each model’s ability to identify patients at risk for NEC. The area under the ROC curves from the final prediction logistic model and the final classification tree were compared using the methods of Hanley and McNeil[152]. Sensitivity and specificity of the two methods were evaluated, as well as the differences between logistic regression and the CART analysis with regards to the assigned predicted probabilities. The CART tree can provide a probability that is interpreted as the prevalence of the outcome at each terminal node, which gives a reasonable approximation of the probability [153].

**Specific Aim 3**

The purpose of Specific Aim 3 is to determine the clustering of cases of NEC within a NICU setting. The selection of records used for this aim has been explained in Specific Aim 1, but included all races and was limited to years 1996-2004, since all three NICUs were involved with the NICHD NRN by this year. The incidence of NEC was defined as the incidence of newly developed NEC per 100 admissions. The incidence was calculated by year, by month within each year, and by hospital site. The incidence density rate was defined as the number of newly developed NEC cases per 1000 patient-days. The rate of NEC at each time point was compared to the overall rate calculated for years 1996-2004 (the baseline rate). Poisson regression was used to investigate the potential for seasonal and secular trends in NEC rates.

Scan statistics have been applied to test for temporal clustering of a variety of adverse health events. These events include the possible clustering of lung or breast cancers [154, 155], adverse events following vaccination trials in children [156], birth defects [157, 158], toxic events for poison control surveillance[159] and biologic terrorism[160]. A modified version of the scan statistic was used to investigate the potential for the occurrence of the clustering of episodes close together in time. First proposed by Naus, this test statistic is the maximum number of cases (n) of interest observed in a time window of length t [161]. It is found by
scanning all intervals of length \( t \) over a given time frame \( T \). Wallenstein later created tables of \( p \)-values to be used for determining statistical significance of the scan test statistic \([162]\). Grimson and Mendelsohn further modified this cluster detection method by basing the scan statistic on the binomial test\([140]\): \( P(R > \text{max} \mid T, N, k) \) where “max” denotes the largest observed frequency of events occurring in the most current \( k \) days of the \( T \)-day period. Formulas proposed by Wallenstein and Neff \([163]\) and later modified by Grimson and Mendelsohn \([164]\) were used for the purpose of testing for possible clustering of cases in this study.

The first issue to address for the assessment of a temporal cluster is the choice of window size, which, if possible, should be predetermined and selected in concordance with periods in which a “cluster” or outbreak are considered \([164]\). For example, carbon monoxide poisoning outbreaks are typically acute, often limited to a few days\([159]\). Therefore, the outbreak window considered in this scenario might be limited to 1-3 days. However, the mechanisms of NEC transmission are unknown, which increases the uncertainty of defining an important cluster window. Reports on the duration of clusters and/or outbreaks of NEC are inconsistent with respect to the time the cluster was observed, ranging from several weeks to several months (Table 3). Many case series reporting NEC outbreaks are from institutions that rarely see cases within their population \([79, 84, 85, 89, 92, 96]\), making it difficult to define a cluster without a baseline or expected rate. For the purpose of this study, the window length necessary was based on subjective data from clinicians focusing on time frames during which cases may be considered as clustering, while considering the median duration between cases (time at which no cases are seen). For each year, a time period \((k)\) of 7 days was selected to scan the time frame \((T)\) of 30 days. The maximum number of cases \((n)\) is counted in each time window and a window with a level of significance \( \leq 0.05 \) is then identified. Because this method requires multiple testing, adjustments for this must be taken into consideration. However, there are no agreed upon methods for adjusting for multiple comparisons during a cluster analysis. This
issue is described in detail later in the Discussion Chapter under *Multiple Comparison Strategies*.

A limitation of this scan test, which is based upon the binomial test, is the assumption that the denominator remains static. Unfortunately, as noted by the daily census, the population within the NICUs on any given day may vary greatly. Therefore, a second method of cluster identification was applied that accounted for variability in the actual number of patient-days contributing to the denominator. For every potential cluster identified through the scan test, the incidence density rate (IDR) of NEC was calculated and compared to an expected rate of NEC. Groups of NEC with rates statistically significantly different than the expected rate were noted. The observed incidence density rate (IDR) of NEC was calculated for the same 7-day window that was identified as statistically significant during the scan test. The expected rate, or background rate, was defined as the number of NEC events for years 1996 – 2004, divided by the total number of patient-days for the same years. This type of comparison investigated the possibility that potential clustering of cases may coincide with an increase in the overall rate of disease. The Fisher’s Exact p-value was used to determine statistical significance.

The combination of statistical significance (p<.05) by the scan test and the IDR comparison defined a *consistent cluster*. The label of *inconsistent cluster* was applied when the group of NEC cases was found to be statistically significant by one method (i.e. scan test) but not statistically significant when comparing the observed IDR with expected IDR. *Sporadic cases* were isolated cases (no other cases within a 7 day window) or cases, that appear temporally close to one another (within 7 days of each other), yet are not considered statistically significant (p≥.05) based on both the scan test and the rate comparison.

**Specific Aim 4**

The purpose of Specific Aim 4 is to examine risk factor differences between sporadic cases and cluster cases of NEC. In order to accomplish this aim, the same potential predictor
variables that were explored as part of the prediction model protocol were included. Differences with regards to clinical characteristics between statistically defined consistent cluster cases (see Specific Aim 3 above), and sporadic cases were investigated as well as inconsistent cluster cases and sporadic cases. Medians and interquartile ranges were reported for continuous variables and differences were tested using the Wilcoxon Rank Sums test. Categorical variables were reported as proportions and differences were tested using chi-square or Fisher's Exact when appropriate (expected cell frequency less than 5).
Chapter 4: Results

Study Cohort Description

After limiting the dataset to the inclusion criteria stated in the Methods, and only including non-Hispanic whites and African Americans, 3816 infants were included in the final analyses; 263 infants acquired NEC. The overall incidence rate of NEC for the study period was 6.96 per 100 VLBW admissions (95% CI 6.2-7.7). These rates appeared to be relatively stable from 1991 to 2005, ranging from 3.8% to 11% (Figure 5) and were not significantly different among hospital NICUs (6.7% vs. 6.5% vs. 7.2%; p=0.83). The proportion of infants with NEC increased with decreasing birth weight (p<.0001 Cochran-Armitage test for trend) (Figure 6). The age at onset of NEC was inversely related to the infant’s weight at birth in a simple linear regression model ($\beta = -0.01339$, p<.0001) which is consistent with previous findings (Figure 7). However, further investigation of this relationship showed that this relationship was strongest among infants who were >1000 g birth weight (Figure 8). The median age of NEC was significantly higher among infants <1000 grams (n=1595) compared to infants >1000 grams (n=2221) birth weight (19 days vs. 11 days; p<.0001). Since the age of NEC onset was reported for the sake of describing the population, and not used in any further analyses, the association between the age of onset and birth weight was not fully explored. The mortality rate due to all causes in this population was 10.5% (95% CI 9.5-11.5); mortality was significantly higher (p<.0001) among infants with NEC (32%, 95% CI 26.6 – 38%) compared to infants without NEC (8.8%, 95% CI 7.9-9.8%). Figure 9 displays the mortality rate among these 2 groups (NEC and non-NEC) over time.

Maternal and infant characteristics and main outcome of the 3435 NICU admissions used to develop the model (derivation sample) and the 381 admissions used to test the model (validation sample) are shown in Tables 5 & 6. The derivation and validation samples were similar with respect to gestational age, birth weight, sex distribution, maternal characteristics and infant clinical characteristics.
Predictors of NEC

Thirty-three clinical characteristics were identified in the literature, through clinical plausibility, or through statistical testing as potential predictors of NEC. The characteristics include maternal characteristics (maternal age, marital status, prenatal care, delivery method, pregnancy induced diabetes, pregnancy induced hypertension, antepartum hemorrhage, PROM, maternal antibiotic exposure, antenatal steroids, gravida, parity, multiple births); and infant clinical characteristics (gender, race, birth weight, gestational age, SGA, Apgar score, prolonged rupture of membranes, mechanical ventilation, infant antibiotic exposure, indomethacin, patent ductus arteriosus, surfactant use, delivery room intubation, respiratory distress in the first 24 hours, oxygen requirements in the first 24 hours, respiratory support required in the first 24 hours, abnormal chest x-ray in the first 24 hours, first enteral feeding age, first full enteral feeding age, high FIO2, and admission temperature).

Unadjusted Results

The unadjusted associations between clinical characteristics and NEC were studied by computing the odds ratio and 95% confidence intervals for each characteristic. Among the 3435 subjects in the derivation sample, there were a total of 231 NEC cases (6.7%; 95% CI 5.9, 7.6) and among the 381 subjects in the validation sample there were a total of 32 NEC cases (8.4%; 95% CI 5.6, 11.2). The proportion of NEC cases in each sample (derivation and validation) were not significantly different (p=0.22).

Maternal/Prenatal Characteristics

The results summarizing the association between maternal and delivery characteristics for the derivation sample of 3435 subjects are seen in Table 7. The odds of NEC were slightly higher among those infants who were not the first live birth [parity > 1 (OR 1.3; 95% CI 0.99, 1.8)]. Mothers who were exposed to antibiotics at admission were slightly more likely to have an infant with NEC (1.3; 95% CI 1.01, 1.7) while mothers who had experienced hypertension at
admission or during delivery were less likely to have an infant with NEC (OR 0.75; 95% CI 0.54, 1.04). There was no difference between groups with regards to the age of the mother, marital status, delivery method, or whether this was a multiple birth (twins, triplets, etc).

Clinical Characteristics of Infant

The results summarizing the clinical characteristics of infants are seen in Table 8. Compared to infants without NEC, those who acquired NEC were more likely to be African American, have lower birth weight, have 5-minute Apgar < 7, be on mechanical ventilation for > 2 days in the first week of life, be exposed to early round of antibiotics, received indomethacin, and have PDA. Infants with NEC were also more likely to require intubation at delivery and in the first 24 hours of life present with respiratory distress, require respiratory support, require oxygen, and have an abnormal chest x-ray. Infants with NEC also had high FIO₂ findings and low temperature at admission compared to infants who did not develop the disease. Infants with NEC were less likely to achieve full enteral feeds by 5 days of life compared to infants without NEC.

Model Development

In a multivariable analysis, clinical characteristics were combined as predictor variables in a logistic regression model predicting NEC (outcome). For each infant in the multivariable analysis, the probability of NEC was calculated from the regression model (predicted probabilities). The reliability, discriminative ability, and validity of the model were assessed as stated in the Methods. Observations with missing data were excluded from the logistic regression analyses, though there were few missing data points (n=15 missing values in the overall logistic regression analyses).

Birth weight was entered into the logistic regression model as a continuous variable, although birth weight was also dichotomized at 1000 grams for additional models. Feeding information was evaluated at the age of receiving the first enteral feeding and the time it took to
achieve full enteral feeds. In this population of VLBW infants, the mean [SD] age of receiving the first enteral was 3.8 days [4.4], while the median age was 2 days; 206 infants were never fed enterally prior to death or discharge from the NICU. The mean age of first enteral feeding among infants with NEC was 4.8 [6.5] days while the mean age of first enteral feeding among infants without NEC was 3.7 [4.2] days (p=0.007). The variable for age at receiving first enteral feeding was dichotomized at ≤ 3 days and > 3 days for the purpose of the logistic regression. Since the purpose of evaluating this feeding variable was to determine if being fed early (first few days of life) was a risk factor for NEC, infants who were not fed enterally (did not have a first enteral feeding date) were classified as receiving first enteral feeding > 3 days.

Based on the results from the bivariable analyses (Tables 7 & 8), 22 clinical characteristics that had p-values < 0.2 in bivariable analyses, were selected for the regression model by stepwise selection of the most significant characteristics, done by using the Akaike information criterion. These variables were marital status, maternal hypertension, PROM, maternal antibiotic exposure, infant parity, multiple birth, race, birth weight, SGA status of infant, 5-minute Apgar score, mechanical ventilation, infant antibiotic exposure, indomethacin, PDA, surfactant use, delivery room intubation, required oxygen in first 24 hours, showed respiratory distress within first 24 hours, required respiratory support in first 24 hours, and had abnormal chest x-ray in first 24 hours, age at first enteral feed, and age at full enteral feed. As a result, 15 variables were dropped from the model (p>0.15): gender, maternal age, mechanical ventilation, SGA, PROM, delivery intubation, early antibiotics, low 5-minute Apgar score, mechanical ventilation, infant antibiotic exposure, indomethacin exposure, clinically significant PDA, surfactant use, delivery room intubation, oxygen required first 24 hours of age, respiratory distress in first 24 hours, need for respiratory support in first 24 hours, delivery method, maternal hypertension, indomethacin treatment, antepartum hemorrhage, antenatal steroids, multiple births, and maternal antibiotics. Four variables (required oxygen in first 24 hours, showed respiratory distress within first 24 hours, required respiratory support in first 24 hours,
and had abnormal chest x-ray in first 24 hours) were considered for the logistic regression analysis, though 8% of the data were missing from these variables (n=298 missing). Two additional variables were not considered (High FIO2 and low temperature at admission), since these variables were collected only for brief periods of time and are missing for 40% and 71% of subjects respectively. Interactions between birth weight and clinical characteristics and race and clinical characteristics were studied in 2 ways. First, a likelihood ratio test on all first 2-way interaction terms was performed. Second, biologically plausible interaction terms were tested.

**Results of Logistic Regression Analysis on the Derivation Sample**

Prior to the addition of any interaction effects, the strongest predictors of NEC, according to their Wald statistics, were the continuous variable birth weight ($\chi^2 = 47.85$) and PDA (as a categorical variable) ($\chi^2 = 25.3$). The derivation logistic regression model contained the following variables, selected through stepwise methods: birth weight, African American race, achieving full enteral feeds by 5 days, PDA, and parity of child. The regression coefficients are shown in Table 9. Two-way interaction terms between birth weight and other covariates, and race and other covariates, were tested in the model, resulting in significant interaction between birth weight and parity of infant (Table 10).

The final model created from the derivation sample is shown in Table 10. In the multivariable model, infants had a greater likelihood of acquiring NEC if they were African American (OR 1.40), and had PDA that was not treated (OR 1.73), or had a PDA that was surgically treated (OR 2.82). Infants had a lesser likelihood of acquiring NEC if they managed to achieve full enteral feeds within the first 5 days of life (OR 0.25). In terms of risk, one can also state that infants who were unable to achieve full enteral feeds within the first 5 days of life were more likely to acquire NEC (OR 4.0). This particular finding, with regards to enteral feeding, may represent the health status of the infant; a healthier infant may be more likely to achieve full enteral feedings early in life. It is interesting that there is a modifying effect of birth
weight with regards to parity. For example, among infants who are not first born (parity > 1), the likelihood of NEC increases by 1.18 (95% CI 1.11, 1.26) for every 100 gram decrease in birth weight (p<0.0001). This effect between birth weight and NEC is not apparent among infants who are first born (OR 1.04 95% CI 0.95, 1.14). This effect modification will be further described later in this chapter.

The apparent calibration of the derivation sample model was satisfactory (Figure 10), without evidence for a poor fit (Hosmer-Lemeshow statistic p=0.27). To study discrimination of the prediction model, the area under the ROC curve (AUC) was calculated. The area under the ROC curve (AUC) is an estimate of the overall predictive or diagnostic performance of the model that incorporates both specificity and sensitivity of the model. Thus, this test of model performance assesses the ability of the model simultaneously to exclude infants who will not develop NEC and include infants who will develop NEC. The prediction model of the derivation sample had an ROC area of 0.685 (95% CI 0.652 – 0.719), indicating moderate ability to distinguish between infants who will experience an episode of NEC and those who will not.

**Validation Sample**

The performance of the prediction model was tested on the internal validation set (10% of original sample; n=381) (Table 11). Using the same predictor variables that were included in the derivation model, the AUC for the validation sample was slightly improved compared to the derivation sample model (AUC 0.734, 95% CI 0.642– 0.830) (Figure 11). The predictive value (regression coefficients) of most predictors of NEC in the validation sample were similar to that in the original sample (Table 12). Achieving full enteral feeds, however, seemed positively associated with NEC in the validation sample, though the odds ratio was not statistically significant (OR 2.51, 95% CI 0.62 – 10.21). Although this appears to be a serious issue with regards to the validity of the inclusion of this variable in the model, the confidence interval is extremely wide, and this may be a result of an inadequate sample size for the validation sample. The sample size of the validation sample limits the power to detect the same statistically
significant predictors, a limitation that is related to utilizing a 90% derivation/10% validation split. When the original derivation model was applied to the validation sample with the regression coefficients fixed, the sensitivity and specificity was similar. For example, at a 6% probability cutpoint (assuming that infants with 6% or greater predicted risk of NEC will develop the disease), the sensitivity [95% CI] and specificity [95% CI] of the derivation sample was 67% [61-73] and 58% [56-60] respectively. When the coefficients were applied to the validation sample at the same probability cutpoint, the sensitivity and specificity was 72% [56-88] and 58% [53-63] respectively. No statistical difference was seen between the samples with regards to the sensitivities (p=0.70) or specificities (p=0.98). Other validation methods have been employed, such as an external validation of the prediction model, which is considered a more powerful validation method than internal validation methods[123, 127]. Results of the external validation are reported later in this chapter.

The agreement between the predicted probabilities and the observed frequency of NEC for both the derivation and the validation models is shown in Figure 12. The deviations from the ideal line were not statistically significant for either the derivation or validation models (Hosmer-Lemeshow goodness-of-fit test p=0.27 and 0.28 respectively). It appeared that the predicted probabilities of NEC in the higher range were somewhat too high in the validation set. In the validation group, there were 6 cases of NEC among the 190 infants in the lower half of predicted risk, an incidence of 3 cases per 100 (95% CI 0.67, 5.65). There were 13 NEC events among the 41 infants in the top decile of predicted risk, an incidence of 31.7 per 100 (95% CI 17.5, 45.9).

**Combined Model Results**

The model was fitted for the whole population (n=3816); the unadjusted risk factor analyses are presented in Tables 13 and 14 and the regression coefficients are summarized in Table 15. No subjects were excluded from this analysis due to missing values. The final
logistic regression model contained the same variables that were tested in the validation sample model: birth weight, race, achieving full enteral feeds by 5 days, PDA, and parity of child. Infants had a higher likelihood of NEC if they were African American (OR 1.51, 95% CI 1.16, 1.96), had a PDA requiring surgical repair (OR 2.83, 95% CI 1.79, 4.49), or had a PDA that received no treatment (OR 1.99, 95% CI 1.31, 3.01). Infants who had achieved full enteral feeding within the first 5 days of life were less likely to acquire NEC (OR 0.40, 95% CI 0.19, 0.88). It is interesting to note the statistical significant finding of parity in this model, and the interaction between parity and birth weight with regards to likelihood of NEC. Among infants who were not first born (parity > 1), a decreasing birth weight corresponded to an increase in the likelihood of NEC (OR 1.18, 95% CI 1.11, 1.26. This modifying effect of parity on birth weight does not hold true among infants who are first born (parity = 1). Table 16 summarizes results from a logistic regression model without including the interaction term between parity and birth weight. The beta coefficients of this model are consistent with the model containing the interaction term.

The predicted probabilities for the infants with NEC (mean ± standard deviation, 9.6 ± 5.3%) were higher than for the infants without NEC (6.7 ± 4.4%, p<0.0001) (Figure 13). To study discrimination of the prediction model, the area under the ROC curve was calculated. The final risk prediction model had an ROC area of 0.69 (95% CI 0.66 – 0.72), indicating only moderate ability to distinguish between individuals who will be diagnosed with NEC and those who will not, consistent with initial derivation sample findings. Probability cutoffs are displayed on Figure 14 to show the relationship between different probability levels and the sensitivity/specificity of the model. For example, with a probability cutpoint of 8%, 129 NEC events and 2564 non-events were classified correctly (71% correct classification). On the other hand, 974 subjects were misclassified as NEC cases (false positives), and 134 subjects with NEC were misclassified as non-cases (false negatives). A comparison of the full model with all candidate variables with the current, parsimonious model is described later in this chapter.
Stratified models

Additional regression models for NEC and risk factors were developed to further understand how birth weight may be modifying the effect of the risk factors on the outcome. The data were stratified by birth weight (401-999 grams, n=1595; 1000-1500 grams, n=2221) and both bivariable and multivariable analyses were conducted. The data were stratified at 1000 grams (g) based on 2 reasons: 1) a cutoff at 1000 g is a clinical dichotomy representing infants who are extremely low birth weight (< 1000 g) and infants who are very low birth weight (≥1000 g); and 2) a 1000 g cutoff signifies the optimal cutoff point with the best sensitivity-specificity trade off, which is explained in greater detail earlier in the Methods under Dichotomized and Calculated Variables.

The proportion of infants who acquired NEC was significantly higher among infants <1000 g (9.7%, n=154) compared to infants ≥1000 g (4.9%, n=109) (p<.0001). Differences in clinical characteristics between birth weight groups are shown in Tables 17 & 18. Infants < 1000 g acquired NEC at a much later date than infants ≥1000 g (21.8 vs. 14.3 days; p<.0001).

Compared to mothers of infants who were ≥1000 g birth weight, mothers of infants who weighed < 1000 g at birth were more likely to have antepartum hemorrhage (23% vs 17%, p<.0001), more likely to be unmarried (52% vs. 47%, p=0.004), less likely to have pregnancy induced hypertension (24% vs. 27%, p=0.018), and less likely to have the current pregnancy result in a multiple birth (25% vs. 31%, p=0.0003). There appeared to be many differences with regards to infant clinical characteristics between infants who were < 1000 g birth weight and infants ≥1000 g birth weight (Table 18). No significant differences existed with regards to antenatal steroid use, maternal antibiotic exposure during delivery, or the gender of the infant.

Unadjusted associations between NEC and clinical characteristics are shown in Tables 19-20 for infants with birth weight < 1000 g and infants with birth weight 1000-1500 g. There appeared to be differences in the unadjusted risk factors for NEC between infants < 1000g and infants ≥1000g at birth. Among infants who were < 1000 g at birth (n=1595), infants with NEC
compared to non-NEC infants tended to be slightly smaller at birth (746g vs. 778g, p=0.007), more likely to have a surgical PDA (12% vs. 7% p=0.032), require antenatal steroids (81% vs. 70%, p= 0.008), and slightly more likely to require mechanical ventilation for >2 days in the first week of life (62% vs. 55%, p=0.099).

The potential predictors of NEC appeared to be slightly different for infants 1000-1500 g birth weight. In this birth weight category (n=2221), compared to non-NEC infants, mothers whose infants acquired NEC tended < 21 years of age, (33% vs. 23%, p=0.017), and were less likely to have pregnancy induced hypertension (14% vs. 28%, p=0.001). Compared to non-NEC infants, infants who acquired NEC were more likely to be African-American (43% vs. 29%, p=.002), to have had early exposure to antibiotics (41% vs. 30%), less likely to be small for gestational age (11% vs. 25%, p=0.001), and less likely to have achieved full enteral feedings within the first 5 days of life (5% vs. 16%, p=0.001). It is interesting to note that these predictors for the larger infants listed above were not significantly associated with NEC among infants who were < 1000g.

Results of the stratified logistic regression analysis illustrated that, among infants who were < 1000g at birth (n=1595), decreasing birth weight, surgical PDA, parity of the child, and antenatal steroids were significantly related with an increase in the likelihood of NEC (Table 21). However, among infants who were ≥1000g at birth (n=2221), being African-American, having a surgical PDA, having an untreated PDA, and having a mother < 21 years of age were associated with an increased likelihood of NEC (Table 22). Achieving full enteral feeds within 5 days of life and maternal hypertensive disease of pregnancy were both negatively associated with NEC in the multivariable model.

Several interesting findings are apparent with this stratified analysis. The first is the appearance of new predictors in both models that were not considered statistically important in the overall model (n=3816). This includes the use of antenatal steroids as a predictor for NEC among infants who are < 1000g and the variables for maternal hypertension and maternal age...
for infants who are ≥1000g. These variables were not found to be significant in the overall prediction model, or were not selected during the model derivation process. It is possible that birth weight was modifying the effect of certain variables in such a way that they were not statistically significant alone, but would become statistically significant as an interaction term. In order to investigate whether this was true, all 2-way interaction terms would require being entered into all multivariable models, regardless of whether the main effect itself was statistically significant. The following paragraphs discuss these birth weight specific model findings.

In order to understand the modifying effect of hypertension on birth weight as it relates to NEC, the significant interactions between hypertension and birth weight were graphed for the entire sample (n=3816) and for the subset of infants ≥1000 g (n=2221). The predicted logits were plotted for specified values of birth weight (see detailed description in the Methods section). Figure 15 shows a scatter plot of the predicted logits by birth weight at the 5th, 25th, 75th, and 95th percentile for the entire cohort of infants. Although for both groups (hypertensive and non-hypertensive mothers) the probability of an infant with NEC decreases as birth weight increases, there is a cross-over near the 25th percentile (birth weight ~ 800 g). Among hypertensive mothers, the probability of NEC is decreasing with increasing birth weight, and this decrease in probability appears to be more prominent compared to mothers without hypertension.

Another intriguing finding is in the differences in the discriminative abilities of the 2 stratified models (Figure 16). According to the AUC, the prediction model comprised of infants < 1000 g birth weight had poorer ability to discriminate infants who would acquire NEC (AUC=0.61, 95% CI 0.56, 0.66) compared to the prediction model comprised of infants ≥1000 grams birth weight (AUC = 0.73, 95% CI 0.69, 0.78). By stratifying the logistic models, it appears that the one for VLBW infants is better able to predict the likelihood of NEC among infants. For both stratified models, the predicted probabilities for infants with NEC were greater than the predicted probabilities for infants without NEC (Figure 17). Among infants < 1000 g
birth weight, the mean probability for infants with NEC was 11.5 (± 5.3) while the mean probability for infants without NEC was 9.5 (± 4.0) (p<.0001). Among infants 1000-1500 g birth weight, the mean probability for infants with NEC was 9.7 (± 9.1) while the mean probability for infants without NEC was 4.7 (± 4.4) (p<.0001).

Based on the findings from the stratified analysis, the following variables were entered into the overall regression model (reported in Table 16): antenatal steroids (significant predictor among infants < 1000 g), maternal hypertension, and maternal age < 21 years (significant predictors among infants ≥ 1000 g). It is important to note that these three variables were not selected during the derivation of the prediction model as important predictors of NEC. When placed into the final model, antenatal steroids was a statistically significant predictor for NEC (OR=1.44 95% CI 1.06, 1.94), though the interaction term for steroid use and birth weight was only marginally significant (p=0.1). After conducting modeling diagnostics, (i.e. examining observations with very high DFBETAS), the deletion of a few observations (between 5 and 10) substantially altered the values of the parameter estimate for steroid use, so that it was no longer statistically significant. Therefore, the decision was made to not include antenatal steroid use in the overall model (n=3816). However, because it appeared to be an important predictor for infants < 1000 g, and remained significant after conducting model diagnostics for that subset of infants, antenatal steroid exposure remained in the final model for infants < 1000 g birth weight (Table 21).

Maternal age was not statistically significant when included in the overall model (OR=1.19; 95% CI 0.88, 1.62), and when it was entered into the model, the Hosmer-Lemeshow (H-L) goodness of fit chi-square became statistically significant (p=0.037), placing the model fit into question. The interaction term between maternal age and infant birth weight was tested since the discovery of this variable as a potentially important predictor occurred only after stratification by birth weight. This interaction term was not statistically significant (p=0.06) and ultimately the variable for maternal age was not included in the overall model.
Hypertension was also not statistically significant when included in the overall model (OR=0.84; 95% CI 0.61, 1.16), although the area under the ROC curve was increased slightly from 0.689 to 0.693; not a large increase in the ability to predict the onset of NEC. The H-L statistic indicated a fairly good fit (p=0.13). Because maternal hypertension was only found to be an important predictor among infants > 1000 g birth weight, the interaction between hypertension and birth weight was tested and found to be statistically significant (p=0.02). Diagnostics for the model which included this interaction term indicated that the model was not influenced by outliers (DFBETAS and influential plots). After testing the contribution of this variable in the model (-2 loglikelihood and AIC comparisons) the decision to keep both hypertension and the interaction term of hypertension with birth weight was made. Tables 23 and 24 show the results from this model. Among infants whose mother had hypertension during pregnancy, for every 100 gram decrease in birth weight, the likelihood of NEC increases by 1.15. For example, among mothers with hypertension, an infant who weighs 1200 grams at birth will be 1.15 times more likely to develop NEC compared to an infant who weighs 1300 grams at birth.

For a better understanding of the candidate models that have been discussed, and the importance of parity and hypertension in the overall model, Table 25 shows the AICs and the ROC curves for each of the 4 candidate models discussed: Model 1 Birth weight, race, feedings, PDA, parity; Model 2 Birth weight, race, feedings, PDA, parity, parity x birth weight interaction; Model 3 Birth weight, race, feedings, PDA, parity, parity x birth weight interaction, hypertension; Model 4 Birth weight, race, feedings, PDA, parity, parity x birth weight interaction, hypertension, hypertension x birth weight interaction. Because Model 4 had the highest ROC curve (0.69) and the lowest AIC, it was included as the final model on which the external sample was tested.

A full model, containing all of the candidate variables chosen as potential predictors of NEC (a total of 18 predictors with bivariable statistical significance p<0.20 from Tables 13 & 14) yielded an AUC of only 0.696 (95% CI 0.66, 0.73) for the entire sample (n=3816). It appears
that the addition of all potential variables does not add much to the overall predictability of NEC in this sample. The large chi-square p-value (p=0.94) comparing the full model with the final parsimonious model indicates no significant advantage of the additional variables. Comparisons were made of the area under the ROC curves between the final model and a series of simple logistic regression models each containing 1 main effect (Table 26). The AUC for the multivariable logistic regression model was significantly higher for each of the simple model comparisons.

Parity and NEC

In order to understand the unique relationship with parity and risk of NEC in this population of infants, bivariable analysis with parity and other characteristics was conducted. The results of this unadjusted analysis are shown in Tables 27-28. Briefly, mothers who were multiparous (parity>1) were older (27.7 vs. 23.4 yrs, p<.0001), less likely to be single (43% vs. 50%, p<.0001), more likely to have antepartum hemorrhaging (21% vs. 18%, p=0.025), and PROM (17% vs. 14%, p=0.03). Although none of these factors were found to be significantly associated with NEC in either the unadjusted analyses (maternal age, marital status, antepartum hemorrhage; all p>0.1) or in the logistic regression (PROM), when evaluated as a group, they may suggest a complex web of health status indicators for the infant. Infants with parity > 1 (mothers were considered multiparous) tended to have a larger mean birth weight compared to infants who were first born (1068 g vs. 1040 g, p=0.002). There were no differences with regards to most of the infant clinical characteristics between parity groups. This finding may support the importance of the health status of the mother and its role in determining infant risk for NEC. Figure 18 is a graphical representation of the interaction between parity and birth weight as described in the overall logistic model. The predicted logits were plotted by birth weight at the 5th, 25th, 75th, and 95th percentile for the entire cohort of infants.
**Hypertension and NEC**

As with the variable for parity, Tables 29-30 report the unadjusted relationships with hypertension and characteristics that may be risk factors for NEC. Mothers with hypertension were slightly older (27 vs. 25.8 yrs, p<.0001), less likely to have had antepartum hemorrhaging (8% vs. 24% p<.0001), PROM (3% vs. 20%, p<.0001), or antibiotic administration prior to delivery (31% vs. 62%, p<.0001). Infants whose mothers had hypertensive disease at pregnancy were less likely to be male (22% vs. 50%, p=0.001), require antibiotics early (24% vs. 48%, p<.0001), have an Apgar at 5 minutes < 7 (22% vs. 32%, p<.0001), but were more likely to be small for gestational age (48% vs. 14%, p<.0001).

**External Validation**

External validation aims to address the accuracy of a model in patients from a different but plausibly similar population [127]. To determine the generalizability of the derived prediction model to new infants, the original internally validated model was applied to a new data set (external sample, n=39,940). As described in the Methods, the model was applied with the regression coefficients fixed to determine the sensitivity and specificity of the current model on a new but similar population, and the regression coefficients were re-estimated on the new sample to determine if the variables originally chosen were appropriate for the new population. This comparison was performed in order to test the sub-hypothesis of Specific Aim 1 that the multivariable associations of the predictors with the outcome in the external sample would not differ from those in the original internal sample (n=3816). To test this hypothesis, the re-estimated regression coefficients in the external sample were compared to the regression coefficients from the internal sample. The final model, described in Table 24 (also Model #4 in Table 25) was applied to the external sample, as well as the prediction models stratified by birth weight (corresponding to Tables 21 and 22).
The incidence rate of NEC of the external sample was comparable to the rate of the internal sample \( (n=3816) \): 7.1% \( (95\% \text{ CI} \ 6.8, 7.4) \) of infants in the external sample acquired NEC compared to 6.9% \( (95\% \text{ CI} \ 6.1, 7.7) \) \( (p=0.63) \). Tables 31-32 summarize the unadjusted associations between NEC and maternal and infant clinical characteristics. Based on the size of the external sample \( (n=39,940) \), there was ample power to detect very small differences between groups that may not be necessarily considered clinically relevant.

To determine the generalizability of the variables selected in the original prediction model for necrotizing enterocolitis, the final model was applied to the external sample of 32,770 infants who were African American and non-Hispanic White, and the regression coefficients were re-estimated. Figure 19 shows the calibration of the model in the external sample \( (\text{H-L statistic} \ p=0.02) \). The area under the ROC curve \( (\text{AUC}) \) of the external validation sample \( (0.66, 95\% \text{ CI} \ 0.65 - 0.67) \) was similar to that of the internal sample \( (0.69, 95\% \text{ CI} \ 0.66 – 0.72) \). The results of both models (the final model and the external validation model) are shown in Table 33. Although the regression coefficients seemed consistent with the findings from the internal sample model, the variable for parity status was no longer statistically significant. In the internal sample, the interaction between birth weight and parity was statistically significant; parity status (being the first live born) modified the effect of birth weight on the likelihood of acquiring NEC. As described earlier, among infants who were not the first born child, for every 100 gram decrease in birth weight the odds of acquiring NEC increased by 1.15 \( (95\% \text{ CI} \ 1.08, 1.23; p<.0001) \). However, this modifying effect of parity on the relationship between birth weight and NEC does not appear in the external validation sample \( (p=0.98) \). In addition, the main effect of parity itself is not significant in the logistic regression model \( (p=0.74) \). This finding is suggestive of the variable being a site-specific variable, important to the population of infants seen in the Cincinnati NICUs, but not necessarily helpful in predicting NEC among infants outside of this region.
It is important to note that in the internal sample (n=3861), hypertension was not considered an important predictor until after the model was stratified by birth weight; only then was hypertension found to be statistically significant and only among infants who were ≥ 1000 g at birth. In the total model (n=3861) hypertension is a statistically significant predictor when placed as an interaction term with birth weight (p=0.02). In the external validation sample, hypertension appears to remain statistically significant as a main effect (p<.0001) and as an interaction term (p<.0001). The regression coefficients appeared consistent with the original internal sample. When the variable representing parity was removed from the model, the AUC remained the same (Table 34).

The birth weight stratified models (401-999 g and 1000-1500 g), were also applied to the external data, and the regression coefficients were re-estimated. For infants 401-999 g, parity was consistent with the overall model and was not considered an important significant predictor. The other predictors, birth weight, surgical PDA, and antenatal steroid use were all consistent with the original internal model, though the effect size of each predictor was not as large for the external sample (Table 35). The AUC for this model, after applying it to an external of 17,220 infants, was lower (0.56; 95% CI 0.55, 0.58) than when applied to the internal sample (0.61; 95% CI 0.56, 0.66). According to the smaller area, it appears that the ability to discriminate infants who will acquire NEC and infants who will not for infants 401-999 g at birth is not much higher than what could be due to chance. This finding, though not entirely surprising, is interesting and speaks to the difficulty of identifying high risk infants, at least among the extremely low birth weight infants. Among infants 1000-1500 g, results from the external validation of the model were consistent with the original results reported on the internal model, though the effect sizes were smaller in the external model findings (Table 36). The AUC was lower in the external model (0.66; 95% CI 0.64, 0.68) compared to the internal validation model (0.73; 95% CI 0.68, 0.78), but this is not surprising. Models tend to have better discrimination abilities on the populations on which they were created[150].
The current internally validated models (overall and stratified) were applied to the external sample with the regression coefficients fixed. The regression coefficients as reported in Table 24 (original sample) and Tables 21 & 22 (stratified samples) were applied to the external sample in order to determine if the current model, when applied to another population, would yield similar probabilities. The sensitivities and specificities at different probability cut points were examined and compared with the same probability cut points from the original model sample (internal sample). The probability cutoffs utilized, which were chosen at the optimal sensitivity/specificity trade-off for each of the models, are listed in Table 37. For example, the model pertaining to 401-1500 g infants utilized a probability cutoff of 6% (assumed that infants who are at a 6% or greater predicted risk of NEC will develop the disease). The sensitivity and specificity of the each of the models applied to the external sample were slightly lower than the original internal sample models (Table 37). As stated earlier, models tend to perform better on the population on which they were created. When the regression coefficients were re-estimated, the sensitivities and specificities of the external samples improved slightly.

**NEC or Death**

Because infants who died before acquiring NEC might influence the model’s ability to predict NEC, the effect of death was evaluated. A sub-analysis was conducted, removing 314 infants who died but did not acquire NEC and the logistic regression model was refitted. The parameter estimates for this sub-analysis remained consistent with the original analyses of all 3816 infants, while the area under the ROC curve was slightly improved from 0.694 to 0.718. In fact, recreating the model utilizing the same stepwise selection procedure and candidate variables as described earlier led to the same predictors for NEC with the exception of one additional variable for feeding (receiving the first enteral feeding within the first 3 days of life).
These findings, consistent with the overall prediction model for NEC, indicate that the original validated model was not influenced by the subset of infants that died prior to acquiring NEC.

**Recursive Partitioning**

The purpose of Specific Aim 2 was to create a classification tree utilizing the same candidate variables that were considered for the logistic regression model for Specific Aim 1 above. The same significant risk factors identified for the development of the logistic model were considered for entry into the classification and regression tree analysis (CART), except for birth weight which was entered as a categorical variable (weight < 1000 and weight ≥ 1000). An optimal classification tree using 10-fold validation is shown in Figure 20. The final tree, created using the 3816 subjects, made use of 9 different predictor variables and had 12 terminal nodes. For each node in the tree, the numbers of NEC and non-NEC cases and the variable used to split the parent node are displayed. The percentages displayed in parentheses represent the risk of NEC among those who eventually reached this node. The terminal nodes (the square nodes) have been classified as either high risk nodes or lower risk nodes. For example, infants who are in a high risk terminal node may be misclassified as having NEC (false positives). In contrast, infants who are in lower risk nodes may be misclassified as not having NEC (false negatives). This tree successfully classified 210 (80%) of the 263 NEC cases and 1821 (51%) of the 3553 of the non-NEC cases, giving an overall correct classification rate of 53% for this model (2031 correctly classified of 3816 subjects). The results from the cross-validation of the tree model indicated that 69% of NEC cases were correctly classified and 53% of non-NEC cases were correctly classified for an overall correct classification rate of 54%. The first variable selected for splitting was birth weight.

*Birth weight < 1000 grams*

Among infants who weighed <1000 grams at birth (n=1595), infants who were administered antenatal steroids were classified as high risk (terminal node 10; 11% of infants in
this node were NEC events). Infants who had not received antenatal steroids were further split according to parity. Infants who were first born were in a lower risk terminal node; 3% of infants in this node developed NEC. Infants who were not first born were placed in a high risk node: 9% of infants in this node developed NEC. Results from this stratified group of infants (<1000 g at birth) are consistent with results from the stratified logistic regression with the exception of PDA, which was not considered as an important variable for the recursive partitioning tree.

Birth weight 1000-1500 grams

Among infants who weighed 1000-1500 grams at birth (n=2221), a PDA diagnosis was the most significant split: no clinical PDA and PDA treated with indomethacin only were split to the left, and PDA untreated and surgical PDA were split to the right. Infants who had a surgical or untreated PDA were considered as high risk; 16% of infants in this terminal node had NEC. The next significant split appeared to be achieving full enteral feeds within the first 5 days of life (dichotomous split). Infants who were able to achieve full feeds within 5 days of life had a low rate of NEC (1% in the terminal node). Among infants who did not achieve full feeds, the next split was race. African Americans were considered at a higher risk for NEC than non-Hispanic Whites. Among African Americans, those that did not receive antenatal steroids were classified as low risk of NEC with 4% in the terminal node acquiring the disease. The final split on this branch was maternal hypertension: infants whose mothers had hypertension were classified at a lower risk of NEC (3% in this terminal node) while infants of mothers who did not have hypertension were classified as high risk (10% of this node).

Non-Hispanic Whites had a different set of risk factors. Among this subset of infants, SGA appeared to be the first important split, with infants who were born small for gestational age classified as low risk (<1% of this terminal node had the disease). Otherwise, the next split for infants who were not SGA was intubation at delivery. Infants not requiring intubation at the time of delivery were low risk (4% of this terminal node with NEC), and infants who required
intubation were split one more time into 2 terminal nodes. Infants exposed to antenatal steroids were classified low risk (4% with NEC in this node) and infants who were not exposed to antenatal steroids were classified as high risk (11% with NEC in this node).

In summary, infants were classified as **low risk** if they were stratified into the following terminal nodes (t-nodes): Achieved full feeds by 5 days (t-node 2), African American infants not exposed to antenatal steroids (t-node 3), and African American infants with maternal hypertension (t-node 4). Among Non Hispanic White infants, low risk nodes included small for gestational age (t-node 6), no delivery intubation (t-node 7), and exposure to antenatal steroids (t-node 8). Infants were classified as **high risk** if they were stratified into the following terminal nodes: had untreated or surgically treated PDA (t-node 1), African American infants with no maternal hypertension (t-node 5), and Non Hispanic White infants with no antenatal steroid exposure (t-node 9).

**Recursive Partitioning vs. Logistic Regression Modeling**

The results generated by recursive partitioning (tree analysis) were compared with the logistic regression model created in the first specific aim. The AUCs of both models were compared using the methods described by Hanley and McNeil for AUC comparisons [152]. The decision tree had an AUC of 0.68 (95% CI 0.64 - 0.72), which was comparable to the AUC of the logistic regression model (Model 4) of 0.69 (95% CI 0.66 – 0.73). There was no statistically significant difference in the discriminative abilities of the two methods (p=0.87). The variables chosen for the recursive partitioning model were slightly different than those chosen in the regression model. Antenatal steroid use was not statistically important in the overall logistic regression model (n=3816), although it was an important predictor in the tree analysis. However, when the logistic regression model was stratified by birth weight, antenatal steroid use became a significant predictor of NEC among infants who were 401-999 grams at birth, which is consistent with the results of the tree analysis.
Although both models have similar predictive abilities, the sensitivity and specificity appear to differ significantly between CART and Logistic (Table 38). The sensitivity of the CART model (ability to correctly classify NEC cases) was 80% (95% CI 75.2, 84.8). When using a NEC probability cut-point of 6%, the sensitivity of the logistic model was 69% (95% CI 63.4, 74.6). For example, if a logistic regression model was used to help determine inclusion/exclusion criteria for a clinical trial, and a cutoff of 6% was used, then infants who had at least a 6% probability of acquiring NEC would be included. These 2 sensitivities (80% and 69%) were significantly different from each other ($\chi^2$ test for proportions= 8.38, p= 0.004) and could be considered clinically significant if the overall objective of the predictive model is to have a low false negative rate. However, the specificity of the logistic model was higher than that of the CART model (57% vs. 51% respectively), indicating a slight improvement of logistic regression to appropriate classify infants who do not have NEC as not having the disease. The predicted probabilities according to the logistic regression model were compared to the predicted probabilities according to the CART model using the one-sample t test. There was no difference in the predicted probabilities for NEC for individual infants between the two methods (p=0.73).

**Regional Temporal Clustering of Necrotizing Enterocolitis**

The purpose of Specific Aim 3 is to determine the temporal clustering of NEC within a geographic region among very low birth weight infants. Because several studies have reported seeing temporal “outbreaks” of NEC occurring in a region[85, 86, 94, 96], we were interested in determining if cases had occurred as a temporal cluster. As described in the Methods section, definitions were developed for consistent clusters and inconsistent clusters. A consistent cluster was defined as a group of cases occurring within a 7-day period that meet statistical significance (p<0.05) on two methods of testing (modified scan test and observed vs. expected incidence density rate). An inconsistent cluster was defined as a group of cases occurring
within a 7 day period that meet statistical significance (p<0.05) based on one of the two methods of testing. Cases that occurred alone and/or did not meet the criteria for either cluster definition were considered sporadic cases of NEC. Statistically significant clusters that overlap (the end of one cluster overlaps with the beginning of another one) were considered as one cluster.

Between 1996 and 2004, 3070 very low birth weight infants were included in these analyses, with 213 cases of NEC, for an overall incidence of 6.94% (95% CI 6.06-7.90). The median number of cases per year was 25, ranging from 15 to 31 cases; the median number of NEC cases per month was 2, ranging from 0 to 8 cases in any given month. The median duration (range) between NEC events was 11 days (0-111 days). A total of 165,735 patient-days were included for an incidence density rate of 1.29 per 1000 patient-days (95% CI 1.11, 1.45). Although there appeared to be a slight overall increasing trend in NEC cases over the study period (17 cases in 1996 to 26 cases in 2004), the incidence of NEC did not appear to be increasing over time (Figure 21). Seasonal occurrence of NEC was examined through Poisson regression (Figure 22). Although the rates of NEC appeared to be slightly higher during the spring months (March, April, May: 1.39/1000 patient-days) and Autumn months (September, October, November: 1.39/1000 patient-days), no statistical difference among rates existed (p > 0.1).

The expected, or background, incidence density rate of NEC was calculated based on years 1996-2004 and applied to observed 7-day incidence rates. One alternative method of determining expected rates of NEC would be to calculate rates based on data from previous years (1996 -2000, for example) and apply this rate to more current study years (2001-2004). Investigation of the incidence rates for NEC during both time periods (1996-2000 and 2001-2004), led to the conclusion that there was no overall difference in the incidence of NEC (Figure 23), and therefore the background rate for all years was applied. Between 1996 and 2000, 114 cases of NEC occurred among 1639 VLBW infants for an overall proportion of infants with NEC
of 6.96% (95% CI 5.80-8.27). With a total of 87,843 patient-days, the incidence density rate of NEC during this time period was 1.30/1000 patient-days (95% CI 1.07-1.56 per 1000 patient-days). Between 2001 and 2004, 99 cases of NEC occurred among 1431 VLBW infants for an overall proportion of infants with NEC of 6.92% (95% CI 5.66-8.36). With a total of 77,829 patient-days, the incidence density rate of NEC during this time period was 1.27/1000 patient-days (95% CI 1.03-1.55 per 1000 patient-days). There was no difference between time periods (1996-2000 and 2001-2004) with regards to incidence density rate ($\chi^2 = 0.02$, p=0.88) or the proportion of infants with NEC ($\chi^2 = 0.000$, p=0.98).

**Modified Scan Test Results**

Results of the cluster detection analysis for years 1996-2004 are summarized in Table 39. Using the *scan test alone* for detecting potential clustering of NEC cases, a total of 48 statistically significant (p<0.05) potential clusters were identified; an additional cluster (Cluster ID 45 in Table 39) was marginally significant (p=0.092) and remained in the analysis. The total number of potential statistical clusters per year ranged from 4 to 8 with the cluster event sizes ranging from two to five cases; 27 of these potential NEC clusters comprised of only 2 NEC events. The remaining 83 cases were considered sporadic according to this testing method. The 27 potential clusters of 2 events, as well as all other potential clusters, are further evaluated later in this chapter. Although no correction for multiple comparisons was applied to these results, this important issue is addressed in detail in the Discussion.

Among the 49 groups of NEC cases identified as potential cluster events, 8 consisted of overlapping 7-day windows. These clusters are identified on Table 39 as Cluster ID 4, 17, 23, 25, 26, 30, 32, 34, and 38. For the purpose of this analysis, these overlapping periods were considered as one continuous clustering of NEC events.
Observed vs. Expected Results

Once temporal grouping of NEC events were identified as potential statistical clusters by the modified scan test described above, a 2\textsuperscript{nd} statistical method was applied in order to account for differences in the actual incidence rates of NEC. As described in the Methods, the 2\textsuperscript{nd} statistical method relied on an underlying knowledge of the expected, or background, incidence density rate (IDR) of NEC for any given 7-day period. The background rate for NEC was used as the expected rate. Since no differences in the rates of NEC appeared to exist between years 1996-2000 and years 2001-2004 (Figure 23), an overall baseline rate of 1.29 per 1000 patient-days was compared to observed rates of NEC. The Fisher’s exact unadjusted p-value was reported.

Of the original 49 potential clusters identified through the scan test, 17 (35\%) were defined as consistent clusters (p-values < 0.05 for both methods), comprising of 29\% (n=61) of the 213 cases that occurred between 1996 and 2004. The median number of consistent clusters each year was 1 (range 0 in 1997 to 5 in 2001); the median case size of consistent clusters was 3 (range 3 to 5). The median number of consistent cluster cases (cases that occurred during an identified temporal clustering of events) each year was 5, ranging from 0 cases in 1997 to 18 cases in 2001. Year 2001 had the highest proportion of cases that were defined as clusters, with 64\% (n=18) of 28 cases defined as consistent clusters. Figure 24 illustrates the number of consistent and inconsistent clusters by season for each year. There appeared to be no increasing or decreasing trend over the years with the proportion of cases that were considered clusters (Cochrane-Armitage trend test p = 0.07) (Figure 25). Although 35\% of the consistent clusters (n=6) occurred during the spring months, there was no statistical difference in the proportion of clusters by season (chi-square test for equal proportions p=.77). It is interesting to note that although both the number of clusters and the number of cluster-cases was the highest during spring months (6 clusters and 19 cases), the average size of each cluster appeared to be slightly higher during the winter compared to the spring (4 cases per
cluster vs. 3.2 cases per cluster respectively) (Figure 26). As stated earlier, the overall rate of NEC during either the winter or spring months was not different than other months. The median incidence density rate (IDR) of NEC among identified consistent clusters was 7.53 per 1000 patient-days (range 6.4 to 10.5), which was expectedly higher than the IDR (5.17 per 1000 patient-days; range 3.8 to 6.9) of NEC among the identified inconsistent clusters (p<.0001).

Of the 32 inconsistent clusters, 27 were comprised of only two cases of NEC. The remaining 5 inconsistent temporal clusters included 3 cases of NEC in each. It is important to note that all of the statistically significant clusters identified by the scan test method alone, which contained only 2 events, were later defined as an inconsistent cluster when comparing the observed incidence rates with the background rate. This finding may speak to the importance of using 2 statistical methods for identifying a cluster. Because the scan test is based on the binomial test, results for the most recent time window (in this case 7 days) are dependent upon the number of NEC events in the previous 30 day window. Therefore, if no other cases had occurred in the previous window of importance, then an increase of only 2 NEC cases will become statistically relevant. A second level screen for potential temporal clustering of cases can help decrease the number of potential clusters identified based on a previous experience.

Also, it may be noteworthy to report that one additional series of 3 NEC cases (Cluster ID 45) was found to be only marginally significant by the scan method (p=0.092), but had an IDR (5.79/1000 patient-days) that was significantly higher than the background rate (p=0.031). This result is another example of the importance of applying 2 methods for defining statistically significant clusters. The inconsistent cluster described above occurred in February, 2004 (Figure 27). An earlier cluster of 5 NEC events was identified (Cluster ID 44) which occurred during the 30-day scan window. Therefore, the probability of seeing 3 cases in a 7-day period given that 8 cases total have occurred during the 30 day period is higher than 0.05 [P(R≥3 | 30, 8, 7) = 0.092]. Without comparing the observed IDR with an expected IDR, one might draw the
conclusion that no clustering is occurring and ignore a potentially clinically significant grouping of disease events.

**Temporal Clustering by Hospital**

Although the rates of NEC did not differ among the 3 NICU sites (6.7% vs. 6.5% vs. 7.2%; p=0.83), temporal clustering of cases appeared to have occurred at 2 hospitals; NICU A (rate 6.7%) and NICU C (rate 7.2%). The following results will focus on NICUs A and C.

Between 1996 and 2004, NICU A reported 81 cases of NEC in 61,630 patient-days, for an incidence density rate (IDR) of 1.31 per 1000 patient-days (95% CI 1.02 – 1.60). NICU C reported 122 cases of NEC in 88,070 patient-days, for an IDR of 1.39 per 1000 patient-days (95% CI 1.14 – 1.64). There was no difference in rates between the two sites ($\chi^2 = 0.13$, p=0.71).

Table 40 summarizes the results of cluster analysis for all years (1996-2004). Because the denominator for the IDR is lower when calculated by hospital (as opposed to by region as earlier described), the IDRs are higher than the originally reported rates. Therefore, in order to lower the rate of Type I error that may occur when comparing the observed with expected rates of NEC, the significance level was set at $\alpha = 0.01$.

A total of 41 potential site-specific clusters were identified among the two NICUs mentioned above (NICU A and C) for the years 1996 through 2004; 12 (29%) were defined as consistent with a median of 1 consistent cluster per year (range 0 - 4). The median size of these site-specific clusters was 3, with the range being 3 to 4 NEC events. All potential clusters (based on the scan method alone) consisting of only 2 events all were defined as inconsistent clusters. The median IDR for consistent site-specific clusters was 15.3 per 1000 patient-days (range 9.95 to 26.6 per 1000 patient-days). This was significantly larger (p<.0001) than the median IDR for the grouping of NEC events that were identified as inconsistent clusters (10.15 per 1000 patient-days: range 6.06 to 16.95).
Cluster vs. Sporadic cases

The purpose of Specific Aim 4 is to examine risk factor differences in cases that are considered cluster cases and cases considered sporadic cases. The definition of a consistent cluster was applied for the purpose of this analysis. In order to examine the same risk factors that were used for Specific Aims 1 and 2, the defined cases (cluster or sporadic) were matched with the original data set used to create the prediction model, and the risk factors were compared between the consistent clusters (n=57) with sporadic (n=83) and inconsistent clusters (n=62) with sporadic. Because the earlier modeling strategies had in place some exclusion criteria with regards to age and heart or gut malformations and races other than African American and White, the number of subjects included for this aim is slightly different than reported in Aim 3 (202 vs. 213 respectively; 11 subjects excluded).

The clinical characteristics of the consistent cluster, inconsistent cluster, and sporadic groups are presented in Table 41. The proportions of cases with maternal hypertension, PROM, maternal antibiotic exposure at admission and antenatal steroid use was similar in all groups. Although it wasn’t statistically significant, it appears that the proportion of consistent cluster cases that had antepartum hemorrhage might be slightly lower than the proportion of sporadic cases (18% vs. 27% p=0.21). The same thing can be said for parity: mothers in both the consistent group (75%) and the inconsistent group (76%) were more likely to be multiparous compared to the sporadic group (66%), although neither comparison was statistically significant (p=0.24 and 0.21 respectively). It is interesting that the median maternal age is significantly higher for the consistent (26 [interquartile range (IQR) 22-32]; p=0.03) and inconsistent (28 [IQR 22-32]; p=0.027) than the sporadic group (24 [IQR 20-32]).

Infants in the consistent group appeared to be more likely (though marginally significant) to have been exposed to an early round of antibiotics compared to the sporadic group (60% vs. 47%; p=0.14). A higher proportion of infants in the consistent group, compared to the sporadic group, also seemed to require delivery room intubation (54% vs. 41% p=0.12), and to have an
abnormal chest x-ray at 24 hours of life (89% vs. 76%; p=0.04). This group had a slightly higher rate of late onset sepsis compared to the sporadic group, though once again, this was not quite statistically significant (47% vs. 36%; p=0.18). It is possible that a combination of these variables may be a marker for a “weaker” infant, making the infant more susceptible to any environmental pathogens that may be attributed to any temporal clustering of NEC. This study was not powered to address this particular hypothesis, and therefore it is difficult to make any real inferences about differences by groups. However, the results of this specific aim may help shape future hypotheses about this population and about the clustering of NEC events over time.
Chapter 5: Discussion

The overall goal of this study was to explore the clinical and temporal relationships characteristics in relation to risk of necrotizing enterocolitis disease in infants in neonatal intensive care units, and better understand how these relationships may be used in the clinical setting to predict the risk of NEC in this population. Despite the high prevalence of NEC in the NICU setting, there is a lack of epidemiologic research defining both factors that are predictive of this disease and clustering of disease events. This research expands our understanding of the many important predictors of NEC, offers a statistical opportunity to easily monitor potential temporal outbreaks of disease, and has potential for use in clinical settings. With the many remarkable advances that have been made in neonatal medicine, younger premature infants are now surviving well beyond the neonatal period. Unfortunately, these advancements have not been able to decrease the incidence of NEC, but have managed to help infants to survive long enough to acquire it.

Predicting Necrotizing Enterocolitis

As a result of this work, we were able to create and validate a prediction model for necrotizing enterocolitis (NEC) among very low birth weight infants that had moderately fair accuracy (area under the ROC curve of 0.67 after external validation) utilizing 6 factors: birth weight, race, patent ductus arteriosus, enteral feeding, and maternal hypertension. Infant parity appeared only to be important among infants who were part of the internal sample model. Further stratification of the cohort into infants < 1000 grams at birth and infants 1000-1500 grams revealed differences in both the risk factors that predict NEC and differences in the ability to predict NEC. This finding speaks to a potential reason for difficulties clinicians have had in the past in identifying infants who are at high risk for NEC based on the inconsistent list of risk factors that are found in the literature.
Previous studies have reported on risk factors for NEC, but few were able to report on the multiplicative effects these variables have on the disease, and fewer discussed differences with birth weight categories. *This study is the first to show a validated model for predicting risk of NEC in very low birth weight infants (< 1500 g)*. Only a few studies to date have proposed differences in perinatal risk factors for NEC based on gestational age or birth weight. Palmer and colleagues conducted a multi-center case-control study specifically looking at birth weight-specific risk factors for the disease [17]. Investigators stratified infants into 2 birth weight categories (< 1500 g and ≥ 1500 g) and matched infants by duration of hospital stay. Infants who were < 1500 g (n=86) had a slightly different set of risk factors compared to the ≥ 1500 gram infants (n=73). Respiratory distress and umbilical catheterization had the strongest association with increased risk of NEC (RR 19.5 [lower 95%CI 2.4] and RR 18.1 [lower 95%CI 1.7] respectively). Other significant factors included mechanical ventilation, receiving continuous positive airway pressure, and recurrent apnea, all indicative of respiratory problems. Our findings differed, as we did not find associations with regards to respiratory distress-like conditions, and the number of umbilical catheterizations was small (n=547) and not associated with NEC (p=0.68). Palmer et al controlled for birth weight only in the logistic regression analysis, and included only one main effect with each model. Beeby et al found risk factors for NEC only among infants between 30 and 36 weeks gestation, but none for infants who were < 30 weeks [45]. Luig et al also found risk factor differences for infants 24-27 weeks compared to infants 28-31 weeks [43]. In a crude analysis, placental abruption and surgical PDA were the only significant risk factors for preterm infants < 28 weeks, though no risk factor appeared to remain significant in multivariable models for this sub-group. Overall, this extremely premature group had a number of risk factors that made detecting differences between cases and non-cases difficult. Among infants between 28 and 31 weeks gestation, however, findings indicated lower birth weight, younger maternal age, maternal hypertension and PDA as important risk factors for NEC. Results from our prediction models are consistent with the work of previous
investigators. Our study has shown that there exists a different set of risk factors for infants who are < 1000 g at birth than infants who are 1000-1500 g at birth. NEC among infants 1000-1500 g involves additional risk factors other than prematurity, whereas for more premature infants (< 1000 g), additional risk factors do not seem contribute to the risk of NEC. This finding may help explain the poor predictability of logistic models in identifying infants who are at high risk of NEC, particularly among infants < 1000 g at birth.

Our finding of hypertensive disease of pregnancy as protective against NEC, in both bivariable and multivariable analyses, has not been explained fully in the literature. Luig et al had similar findings among infants 28-31 weeks gestation (OR 0.57, 95% CI .35, .92) and suggested that infants delivered after maternal hypertension may have been affected postnatally by cautious feeding introduction and advancement[43]. It could also be hypothesized that infants of mothers with hypertensive disease of pregnancy would be delivered for reasons other than those related to the illness of the infant. In our study, mothers with hypertension were less likely to have prolonged rupture of membranes, less likely to have received antibiotics prior to delivery and were slightly older compared to mothers without hypertension. Postnatally, infants whose mothers had hypertension were also less likely to have required an early course of antibiotics. These factors all may be markers of the illness of the child; mothers who deliver for reasons related to health of the mother (i.e. hypertensive disease of pregnancy) may be delivering healthier infants compared to mothers who are delivering for reasons that may affect the health of the infant (i.e. maternal infection that might affect the placenta). Infants of mothers with hypertension tended to be slightly larger at birth with slightly higher gestational ages. Both of these factors have been known to be inversely associated with the risk of NEC.

The association between parity and NEC did not hold in the external prediction model, but was significant in the original internal sample, and appeared to be a stable predictor in this population, based on the diagnostics run on the logistic models. Among the Cincinnati cohort,
the parity of the infant may be a marker for a risk factor that is not known or is not measured.
Parity has been associated with an increased risk in sudden infant death syndrome among term
infants[165, 166]. Parity may also have a large influence on birth outcome. It has been
reported that among singleton births, infants born to younger primiparous women are generally
at lowest risk of poor birth outcome. Conversely multiparous women are at highest risk of
adverse pregnancy outcomes[167-172].

The relationship reported in these analyses between antenatal steroids and risk of NEC
is not entirely surprising, though somewhat complex. The logistic prediction model was unable
to pick up steroid use as a significant main effect during the derivation of the model, however it
did appear to be related to an increased likelihood among infants who are <1000 g at birth and
was further found to be important in the classification and regression tree (CART) analysis. This
finding speaks to the complexity of the NEC causal pathway. Several studies have shown
contradictory findings with antenatal steroid use. Guthrie et al found an association with
antenatal glucocorticoids and an increased incidence of NEC among 14,878 infants 23 to 34
weeks gestation [42]. Similar were findings by Kamitsuka et al in 2000, who reported a
marginally significant increased risk of NEC among 477 infants weighing 1250-2500 g at birth,
after adjusting for factors such as race, birth weight, and feeding information (OR 2.3; p=0.062)
[173]. Sehdev et al reported an association between the rates of NEC and the duration of
antenatal steroid exposure. An adjusted multivariable analyses of 5 large observational
databases and a total of 11,455 infants (including 9949 infants as part of the NICHD Network)
conducted in 1995 revealed that any steroid exposure was associated with an increased risk of
NEC (OR ranged from 1.3 to 1.8) [174]. All infants included were ≤1750 gram birth weight,
although the range of birth weight varied slightly by database.

Patent ductus arteriosus (PDA) appears to have implications in the pathophysiology of
NEC. It has been postulated that PDA diverts blood from the intestines by shunting. It is most
commonly treated medically with indomethacin, which has both a vasoconstrictive effect and an
association with intestinal perforation [52]. Our findings are consistent with this reported relationship, although when PDA was categorized into untreated, medically treated, and surgically treated, there were differences in risk status. Infants who had an untreated PDA or a surgically treated PDA had a much higher likelihood of NEC compared to infants who did not have PDA or infants who had indomethacin-treated PDA. Not all studies, however, have found a direct association with indomethacin and NEC, particularly when the drug is used for reasons other than PDA closure. Parilla and colleagues performed a case-control study to investigate the possible association of indomethacin tocolysis with neonatal NEC, and found that indomethacin was only associated with an increased risk of NEC when used in conjunction with another tocolytic agent (multivariate regression OR 6.9 95% CI 1.1, 43.6) [35]. However, in our study, we found no evidence that any tocolytic use was associated with NEC in our population of VLBW infants (OR 1.2, 95% CI 0.88, 1.66). Fowlie and Davis reported a systematic review on prophylactic intravenous indomethacin and found no significant difference in rates of NEC among infants who received indomethacin prophylactically versus a placebo (RR = 1.09; 95% CI 0.82, 1.46) [175]. Eleven studies were included when analyzing NEC as the outcome, though the dosing schedules of the drug varied widely. We investigated the relationship between NEC and prophylactic indomethacin in our study, and did find an unadjusted association (OR 1.9; 95% CI 1.2, 3.1). This association did not remain significant in any multivariate analyses. Bellander et al found no differences in the proportion of infants who developed NEC in a case-control study designed to investigate the tolerance to early human milk feeding in relation to indomethacin treatment of PDA among infants < 29 weeks gestation [176]. The investigators noted that based on the low prevalence of NEC (~6%), there was insufficient power to rule out the possible effects of enteral feeding practices on the incidence of NEC. Among extremely premature infants (<27 weeks or <800 g birth weight), early treatment of indomethacin (within the first 48 hours of life) has been shown to be associated with NEC in which intestinal perforation occurred, compared to standard treatment with indomethacin (after
48 hours of life for clinically symptomatic NEC) [177]. This study of indomethacin treatments differed from previous studies in that Fujii et al focused on extremely premature infants, who are at greatest risk of both PDA and NEC. O'Donovan et al found no greater association between indomethacin treatment for significant PDA and NEC than surgical ligation[178] in a large retrospective study of premature infants (n=224). Our findings showed no higher risk for NEC among infants who received indomethacin treatment than infants who did not have clinically significant PDA (OR 1.09 95% CI 0.79, 1.51 of the internal sample). It is important to be aware that most studies, including the large retrospective study, did not conduct multivariable regression to control for other causes or confounders of NEC. This methodological oversight could be one reason, among others, why the literature remains inconsistent.

**Recursive Partitioning for Predicting NEC**

The final tree created by this analysis contained 9 predictors and 12 terminal nodes, with an ability to correctly classify 53% of infants (Figure 20). Predictor variables selected for splitting, in order of importance (higher on the decision tree), for NEC were: birth weight (< 1000 g vs. > 1000 g). Among infants > 1000 g at birth, PDA status was the next split, with surgical/untreated PDA being a high risk node (16% of infants in this node with NEC). Feeding status (ability to achieve full feeds in 5 days) was the next split, with race following in order of importance. The predictors differed between infants who were African American and infants who were non-Hispanic white. The predictors also differed between infants who were < 1000 grams and infants who were > 1000 grams at birth. These variables (race and birth weight) are two potentially important effect modifiers of other predictors for morbidity and mortality among very premature infants, and the modifying effects may be complex that they are not easily noticed with linear regression models.

It is possible that CART analysis might be better suited for understanding the existence of multiple subgroups at risk for NEC and the complex interactions that produce an increased
risk of NEC. The CART analysis conducted for this study displayed a complex relationship between antenatal steroid use and risk for NEC among infants 1000-1500 grams birth weight. As a potential predictor for NEC for this sub-group of infants, antenatal steroid use appeared late in the tree, and only as an important predictor among African American infants who did not achieve full enteral feeds within 5 days of life, who did not have a PDA or did have an indomethacin-treated PDA (Figure 20). Conversely, non-Hispanic White infants exposed to antenatal steroids were classified as low risk for NEC. The overall benefits to the newborn of antenatal steroid treatment often outweigh the risks involved with this treatment. Therefore, the application of multiple statistical methods may be needed in order to understand the underlying complex relationships that are being reported as part of this study, and to provide the best care for preterm newborns.

**Recursive Partitioning vs. Logistic Regression**

The results of this study clearly demonstrates that the ability of a multivariable logistic model to predict NEC among VLBW infants was comparable to a model created from recursive partitioning, also known as classification and regression tree (CART) analysis. Both models (logistic and CART) designated similar birth-weight specific risk factors, both had comparable AUC results (0.68 and 0.69 respectively; p=0.87), and both had similar sensitivities and specificities (Table 38). The CART analysis resulted in similar risk factors as that of the logistic regression prediction model, with the exception of antenatal steroid exposure, small for gestational age, and intubation at delivery among infants 1000-1500 grams at birth. The stratifying nature of the CART analysis is likely the reason why different variables were considered important; these variables seemed to be important only in certain sub-groups of the entire population.

Classification trees and logistic models have been found to have comparable performances [128, 153, 179]. Measures of statistical inference can be obtained through
logistic regression and the magnitudes of the predictors with the outcome can be evaluated. Although estimation of an infant’s individual risk of NEC may be difficult to obtain at the bedside with logistic regression coefficients (as the characteristics of the infant need to be plugged into the equation), there have been ways of making logistic regression results clinically useful. Scoring systems can be developed utilizing the estimated regression coefficients, which would provide a single quantitative value that could be associated to an individual’s risk of NEC. Some investigators have created bedside systems that allow for the clinician to plug in an individual’s characteristics into an electronic source (e.g. palm pilot, interactive web page) which in turn can calculate the estimated probability of the event occurring [180, 181]. The CART models offer a more simplistic approach for determining an infant’s risk, by categorizing an infant as either high risk or low risk of NEC. By forming strata to “adjust” for potential effect modification, CART can be important for generating hypotheses that merit further study.

*Pragmatic use of Prediction Models*

Clinical prediction rules are designed to predict health outcomes and assist clinicians with patient care plans by estimating the probability of a diagnostic outcome. Predictions may be used to classify patients according to prognostic risk, with may be useful for stratifying patients for randomized control trials, studying treatment results, or comparing outcomes of different patient series [182]. For example, when evaluating prevention strategies, clinical trials often enroll subjects that would be considered either “healthy” or not at risk for the disease. Thus, eligible subjects may be exposed to a prevention strategy (e.g. drug) unnecessarily, or they may even shy away from participating due to unknown risk/benefit ratios. Within the NICU setting, prediction models have created widely used severity of illness and mortality scores such as CRIB[117] and SNAP[118] scores.

In the current multivariable model, the independent associations between clinical characteristics and the presence of NEC were assessed. This created prediction rule enables
the clinician to quantify the probability of NEC for any specific VLBW infant. This study had a
variety of strengths according to a set of criteria for clinical prediction rules [101]. The outcome
and predictive variables were clearly defined and clinically sensible. As part of the NICHD NRN,
data is captured in a standardized format. Patient inclusion criteria and all mathematical
techniques were well described.

In order to validate a prediction rule, cross-validation techniques, including the
“jackknife” and “bootstrap” methods can be employed [183]. Bootstrapping can be used in two
ways: 1) as a simple validation of a prediction model, by creating a model through a series of
bootstrap samples (selecting with replacement the full sample size from the original sample)
and evaluating the model on the full original sample; and 2) to estimate the bias (optimism) due
to overfitting the model [184]. For the purpose of this study, the split-sample technique was
used, dividing the data into a derivation set and a validation set. Unlike the previously described
techniques, the split-sample approach does not require resampling of the same population;
rather this approach splits the population into two separate groups. Data splitting has the
advantage of allowing hypothesis tests to be confirmed in the test sample, although it reduces
the sample size for model development, and the chance exists that if the process were repeated
with a different split, different predictive accuracies may be obtained [150]. Although
bootstrapping has an advantage over data splitting as it is useful when the sample size is small,
biases that are introduced through subject selection or data collection procedures are not
eliminated with either method. The “gold-standard” approach to model validation is measuring
the performance of the prediction tool on an independent sample from a different location. This
external validation method was used to externally validate the prediction model for NEC, to
determine if the accuracy of the model was dependent upon practice-specific relationships
between the clinical predictors and outcome. Ultimately, our prediction model continued to
predict NEC among VLBW infants in other populations, though the overall predictability of the
model was not as powerful in external settings.
**Clustering of NEC**

Although necrotizing enterocolitis occurs mostly as sporadic cases, over the past decades, NEC “epidemics” have been reported in the literature, with the number of cases, clinical presentations, and potential causative agent (e.g. pathogen-specific NEC) differing greatly. In 1972, Stein et al reported one of the first outbreaks consisting of 11 cases of NEC in a single nursery in South Africa over a 10-week period [83]. Chappell and Dinner followed this report with another outbreak involving 20 cases over a 3-week period [185]. Only a few studies investigated outbreaks within nurseries that had a base line rate of sporadically occurring disease with periodic increased activity. Rotbart and Levin described a 7-year period in which 193 cases of NEC occurred, with intermittent outbreaks which forced the closure of the nursery[77]. Book et al described 67 cases in a 4-year period, of which 42 cases occurred in a 10 month period [94]. In 1979, Guinan and colleagues reported 3 “epidemics” in 3 different high-risk nurseries, all occurring within the same period[85]. Chaney et al reported on temporal clustering of NEC occurring in 2 different maternity hospitals in Paris in 1982[86], while Faustini et al, in 2004, reported temporal clustering of NEC within 2 different neonatal units simultaneously[96].

Some previously reported outbreaks have occurred in infant populations who are larger than 2000 grams birth weight[84, 85, 91], and many studies reported either very low rates of NEC (i.e. ~11 cases / year) [79, 81, 84, 85, 89, 91, 94] or did not report a background rate[73, 74, 83, 86, 87, 90, 185], leading the reader to assume that perhaps that particular site typically sees no cases of NEC within their institution. Some investigators described both NEC and gastroenteritis cases occurring simultaneously [74, 83, 88, 96, 186]. Because most studies that report NEC outbreaks appear to originate from sites that do not typically report a baseline incidence of NEC, there is likely to be a reporting bias. Institutions that care for premature infants, and have an incidence rate of NEC comparable to the literature (i.e. ~7% among VLBW
infants), may not be reporting periods during which an increase in the number of NEC events occurs close in time. It is probable that institutions that consider NEC to be endemic are not at heightened awareness of this type of clustering.

The results from the current study show that temporal clustering of NEC does occur. According to the definition of a consistent cluster applied to this study, 17 temporal clusters were identified, with no statistical evidence of seasonal variation between 1996 and 2004. It is possible, however, that seasonal variation does exist and this study was not powerful enough to evaluate this hypothesis statistically. Although this study was not designed to address the question of whether the pattern of temporal clustering constitute outbreaks of NEC, our analyses did address the concept of the clustering of events during a specified time period, i.e., 7 days. Since the window length should be predetermined, and meaningful to an “epidemic period” if such a period is definable [164], and with no consensus as to how this period should be defined for NEC, the cluster window of 7 days was used. This window was chosen based on two relevant factors: 1) this duration was less than the median number of days between cases (11 days), and 2) this duration seemed to be clinically and epidemiologically meaningful.

Although this study focused on one particular window of 7 days (k) within a 30 day interval (T), intervals of other lengths may be selected for monitoring. One caveat to altering the intervals is that it is possible to alter both k and T until statistically significant clusters are identified. In order to prevent this type of data mining, the k and T were defined a priori.

**Characteristic differences between cluster and sporadic events**

Using our current definition of a consistent cluster, we did not find any statistically significant differences between the consistent clusters and sporadic cases. This is consistent with current literature. Although previous studies have suggested the existence of perinatal or neonatal risk factors associated with outbreaks of NEC, the majority of these studies were case descriptions and not designed to investigate characteristics associated with the risk of NEC [79,
or found no remarkable characteristics of the infants involved [86, 186]. A few case-control studies conducted to identify risk factors specific to NEC outbreak cases found no new characteristics that are specific to infants who acquire NEC during an outbreak [74, 84, 85, 87, 89, 187], though these investigations compared the outbreak NEC cases to non-NEC controls. No study attempted to compare those cases that occurred in the reported outbreak to cases that occur endemically or sporadically within a specific institution.

We did see some interesting findings, though not statistically significant, that might help shape future study hypotheses. Among infants who acquired NEC, those that did so appeared as part of a consistent cluster to have indications of being "sicker" infants from the start. Compared to infants of sporadic NEC events, those among consistent clusters were more likely to require early use antibiotics, more likely to require intubation at delivery, and more likely to have had an abnormal chest x-ray within the first 24 hours of life. These differences might be markers of infants who have a dysfunctional respiratory system at birth which could lead to a compromise of oxygen levels, ultimately producing a pathway to NEC through intestinal ischemia (see pathogenesis model in Figure 1). Intestinal ischemia has been implicated as one of the three pathological events required to induce NEC [25], though it is unclear whether it is a true cause or an end result. A combination of intestinal ischemia and the immature mucosal barrier could create an environment ideal for the occurrence of NEC through inappropriate bacterial colonization and invasion. However, studies have failed to confirm the association between hypoxic/ischemic factors and NEC development. Circumstances that may compromise oxygen levels of the infant (e.g. perinatal asphyxia, respiratory distress syndrome) and/or reduce mesenteric blood flow (blood flow to the abdominal region) have also been implicated in the pathogenesis of this disease. The infants who were identified in this study to have acquired NEC during a cluster could have been suffering from respiratory-like ailments that made them susceptible to NEC. It could also be hypothesized that it was not the respiratory distress-like condition itself, but that this condition placed these infants in a weaker immune state, making
them more susceptible to gastrointestinal pathogens, another possible cause of NEC. In an attempt to generate these hypotheses, the numbers of children with RSV, flu, and pertussis reported at Cincinnati Children’s Hospital Medical Center were plotted by week for the year 2004, superimposing over the number of NEC cases by week, including the 3 consistent clusters identified in this year (Figure 28). The numbers of the other illnesses were available for year 2004 alone, and reported on the Center’s intranet website, provided by the Infection Control Program Office as regional lab-confirmed illnesses. These numbers were simply considered as a basis for what might be occurring among children in the community. It is imperative for the reader to understand that this figure is not meant to explain the etiology of the clusters that occurred, but rather to give rise to hypotheses behind why NEC clustering might exist. Although no seasonality for NEC rates existed, nor did seasonality appear to occur with NEC clusters, the numbers of consistent clusters for all years were small (17 for all years). If NEC clustering does in fact occur when the rates of community illnesses such as RSV, flu, and/or pertussis are increased, then one would expect to see similar seasonal variation with NEC clustering. Rotavirus-infection rates were unavailable, which could help us understand the potential for a connection between gastrointestinal illnesses in the community and the clustering of NEC. One could stipulate that perhaps something in the community might be making its way into the NICU environment via mother, father, siblings, or healthcare workers. Ultimately, the look at the community illnesses shown in the figure only addresses the fundamental need for a broad thought process in terms of the clustering phenomena. The mechanisms underlying any potentially contagious cause of NEC can only be studied in the context of rigorously-defined cluster events.

In contrast to previous studies that have labeled reports as NEC epidemics or outbreaks, the current study was not designed to prove that epidemics of NEC occur, but rather to show statistically that cases of NEC occur in temporal clusters. A “cluster” is defined as an unusual aggregation, real or perceived, of health events that are grouped together in space, in
time, or both[190]. Clusters of health events, such as chronic diseases and birth defects are often reported to health agencies, identified by either an ongoing surveillance system, but more often identified by concerned groups. Because investigations of the clustering of diseases are often done post hoc, they rarely lead to identifying causation[191], though their importance should not be minimized as they have the potential to generate knowledge. Previous studies have discussed NEC outbreaks within institutions, yet mostly as case studies or reports. *This is the first study to statistically investigate the temporal clustering of NEC events.* By developing a way to help monitor possible true increases in NEC incidence over short periods of time, clinicians may be provided a tool that could guide them with timing the introduction of etiologic investigations and taking infection control measures (i.e. cohorting of infants if infectious etiology is suspected).

The etiology of NEC is most likely multi-factorial, with both infectious and non-infectious factors [1], though the evidence has not fully substantiated an infectious state. Interest in the potential infectious nature of NEC arose with the first reported outbreaks of the disease, though no one specific pathogen has been identified during any outbreak investigation; the most compelling evidence of an infectious etiology lies with the occurrence of outbreaks themselves. The wide variety of organisms that have been implicated in the cause of NEC include *Klebsiella pneumoniae* [84, 85, 90], *Escherichia coli* [81, 192], *Clostridia sp.* [92, 193], rotavirus [74], and coronavirus [86], while some investigations of pathogen-specific causes failed to find any specific organism involved [87-89, 91, 94, 95, 187]. Because the current study was not designed to address pathogen-specific causes of NEC, we were unable to associate any specific pathogen with the disease or with any identified clusters of the disease. It is plausible that the difficulty with linking any specific pathogen as a cause is due to the ability for multiple pathogens to have the same underlying effect on the infant gut, ultimately leading to the gastrointestinal disease. Wide-spread use of antibiotics within the NICU could also make it difficult to detect pathogens when investigating this link.
A number of problems are encountered in the study of clusters. The health events being investigated, in this case NEC, are usually rare, the increase in events tend to be small and may occur over a long period, and information on the population at risk or on expected rates is often unavailable[190]. In the current study, NEC might be considered endemic among VLBW infants in the NICU. Although there were periods of time when no cases occur, the background rate for the disease was approximately 7%, with an incidence density rate of 1.29 per 1000 patient-days; the median duration between cases was 11 days. Therefore, we were able to compare an observed rate of NEC occurring during a suspected cluster to a background, or expected rate. Another issue complicating the investigation of clusters is that some may occur by chance, a concept that is described in greater detail later in this chapter. Our study used a 2-tailed Fisher’s Exact test to test the null hypothesis of no difference between the observed and expected incidence density rate of NEC. The 2-tailed p-values reported for this study were conservative as the alternative hypothesis was that there would be a higher incidence rate of NEC during cluster periods. Regardless of the possibility of “chance” clusters, cluster investigations have historically helped in understanding chronic and infectious disease, providing a means for determining whether an outbreak has occurred. This gain in knowledge is useful regardless of whether the proportion of true clusters is large or small[194].

Several test statistics have been suggested in the literature to statistically approach the investigation of temporal clusters. Ederer, Myers, and Mantel developed a test using a cell-occupancy approach [195]. Naus proposed a scan test for temporal clustering[161], while Weinstock proposed a generalization of the scan test that adjusts for changes in the population at risk[196]. Tango and later Whittemore and Keller proposed a test of temporal clustering based on the distribution of counts in disjoint equal time intervals[197, 198]. Wallenstein and Neff proposed an approximation to the scan statistic that is easy to compute and more accurate than other approximations in the literature[163]. Grimson and Mendelsohn modified this statistic, proposing a more sensitive method for detecting current clusters based on the binomial
distribution[164]. An advantage of using the binomial-like test in lieu of other scan tests is that it specifically tests for clusters in the most recent k days, and therefore may be more powerful than the scan test when monitoring for purposes of immediate intervention[164]. This study has introduced and illustrated a combination of statistical tests (a binomial-like scan test proposed by Grimson and Mendelsohn, and the comparison of observed vs. expected incidence rates with the Fisher’s exact test), that can be easily applied for weekly monitoring of temporal clustering of necrotizing enterocolitis. This study, the first to have addressed the clustering of NEC from a statistical perspective, found evidence of clusters occurring within a 7-day window.

**Multiple comparison strategies**

One of the major statistical issues with utilizing cluster detection methods lies in the Type I and Type II error rate. Type I errors lead to reporting “false” clusters with the possibility of alarming the public unnecessarily. Type II errors lead to missing “true” clusters, and the possibility of harming the public. An appropriate trade-off between Type I and Type II errors should be based on what makes the most clinical and public health sense according to the disease that is being monitored and how it is being monitored. The process of a daily scan, as the one utilized for this study, creates a large number of multiple comparisons, making preset rejection levels for p-values (i.e. Bonferroni adjustments) so small that the power to detect actual clusters not be adequate. Because of the possibility of missing a “real” cluster, using such methods for multiple comparison adjustments are not always recommended[164, 199-201].

Because this cluster analysis was a totally retrospective analysis, we could avoid the multiple comparison issues by simply testing each group of cases once. For example, since we have the ability to “eyeball” the data, we can actually test the cluster once at the last case within an identified group. By utilizing this “loophole” in the definition of multiple testing, the actual statistical testing would not occur until a group of cases were identified. Unfortunately, this
would only work retrospectively, after a group of cases were identified and suspected to have been clustered together. If we are interested in testing each new case as it enters a window of interest with the scan, we could decide to limit the correction for multiple testing to the current identified group of cases. For example, we would begin testing at the 2\textsuperscript{nd} case within the 7-day window. We would scan forward one day, pick up another case, and test again. This would continue until no more events have occurred. In order to correct for multiple testing, we could take a Bonferroni correction, which would use the following formula: 

$$\alpha_B = \frac{\alpha_{FWE}}{c};$$

where $$\alpha_B$$ is the new alpha based on the Bonferroni test, $$\alpha_{FWE}$$ is the family-wise error rate, which is set at .05, and c is the number of comparisons made, in this case it would represent the number of tests (c=number of cases – 1). This Bonferroni correction is considered conservative for this particular analysis, as it assumes independence. However, if we apply this correction method to the “potential” clusters of 3 or more cases (according to the scan p-value on Table 39), 5 of the originally significant groups (at the uncorrected $$\alpha = .05$$ level) would not meet statistical significance at the new level of $$\alpha_B$$. This method of adjustment would correspond to 3 fewer consistent cluster groups (Cluster IDs 30, 32, and 37). Neither of these possible multiple comparison adjustments correct for testing all clusters occurring over all years involved.

Another way to deal with the issue of multiple testing is to consider the false discovery rate (FDR). Instead of controlling the chance for any false positives (as Bonferroni does), the FDR controls for the proportion of false positives, which is determined by the observed p-value distribution. The FDR is defined as the expected proportion of the number of erroneous rejections to the total number of rejections[202, 203]. In contrast to Bonferroni, which seeks to control the chance of even a single false discovery among all the tests performed, the FDR method controls the proportion of errors among those tests whose null hypotheses were rejected. By applying the formulas reported by Benjamini and Hochberg, we would anticipate a false discovery rate of 20\% among the scan p-values. Applying this method to the final 17
“consistent clusters”, we actually have a FDR of 0. Because the FDR is calculated on the ranking of p-values, we may be unable to use any FDR method since we have a number of clusters that are comprised of only 2 cases. From a clinical standpoint, it might make sense to consider all 2-case clusters to be false discoveries, all 3-case clusters to be possible false discoveries, and for clusters of 4 or more cases to be true clusters. This definition for a false discovery, though arbitrary, would allow for a potential FDR of at least 55% among our 49 potential clusters.

Study Limitations

This study is based on retrospective data that is collected for the purpose of the NICHD Neonatal Research Network Generic Database (GDB). Therefore, the predictors that can be evaluated are limited to the data collected for the purpose of the NICHD’s GDB. Information on respiratory distress syndrome, a variable that has been implicated as a potential predictor of NEC in several studies [37-39], is limited to a series of variables that may indicate a variety of respiratory distress-like symptoms (i.e. need for oxygen in first 24 hours of life, required respiratory support first 24 hours). We were able to investigate variables that were available in the GDB that may be related to respiratory distress syndrome and did not find any of them to be significant. The diagnosis of respiratory distress, however, requires a multi-faceted approach, and may vary from one clinician to another. This particular risk factor may not be a feasible predictor for that reason.

One common, yet significant, limitation to this study is the lack of data on the type and amount of milk infants were fed. To date, studies of feeding and NEC are inconsistent, with insufficient evidence to fully guide the feeding management of infants at risk for NEC [204]. Studies on human milk, however, have shown benefit over formula feeding with regards to lower NEC rates [66-68]. McGuire and Anthony conducted a systematic review to determine if enteral feeding with donor human milk reduced the incidence of NEC compared to feeding with formula
milk [68]. Results from a meta-analysis found that infants fed human milk were four times less likely to have confirmed NEC compared to infants fed formula (RR 0.25, 95% CI 0.06, 0.98). No attempt was made to investigate the duration or amount of human milk fed. In an analysis of data collected as part of a randomized controlled trial on glutamine supplementation conducted by the NICHD [205], the amount of human milk fed in the first 14 days of life was associated with a decrease in the risk of NEC or death (hazard ratio 0.87, 95% CI 0.79, 0.97) among extremely low birth weight infants [206]. This finding introduces the potential for a dose-related benefit of human milk fed early in life. We were unable to account for the type of milk infants were fed in our study, which may limit our ability to predict NEC.

A major issue in predictive modeling is the choice of covariables in the model. In the current study, the literature was extensively reviewed for potential predictors. A limitation of this type of review is that most previous studies focused on unadjusted relationships, unable to report potential correlations between predictors. We did attempt to investigate the contribution to NEC of most of these potential predictors, and focused on understanding some of the complex relationships that we saw between factors and risk of NEC. The information on time to event for a variety of variables was unavailable for some of the potential risk factors. Thus, associations with some of the potential risk factors that were investigated could only be described as associations without any real knowledge pertaining to cause and effect. The variables that were ultimately selected for investigation in the multivariate prediction model were chosen as variables that most likely occurred before the onset of NEC. The dates of onset for any co-existing illnesses were not readily available, and therefore couldn’t be used as possible risk factors for NEC. Some circumstances which have been implicated as potential predictors, such as the occurrence of a pneumothorax, could not truly be assessed as predictors since the date of the occurrence was unknown. For variables that dates were known, such as timing of feeds and duration of mechanical ventilation, we chose to focus the risk in what might take place within the first few days to first week of life to limit the number of cases of NEC that would
occur prior to the actual risk factor itself. The model was evaluated after removing those infants who acquired NEC early (first few days), and the results remained consistent with the original findings suggesting that the predictors selected for the model were not influenced by early onset of NEC.

A potential limitation of the use of a scan test chosen for the cluster analysis is the temptation to manipulate the periods or windows of interest to yield statistically significant results. For the purpose of this study, all time periods were set a priori in order to try and reflect what would make clinical sense, as well as exemplify how temporal clustering of NEC can be monitored. As cluster analysis does not include information on exposure, statistical clustering does not prove that the cases have clustered due to some common risk factor or exposure [159].

As stated earlier, the possibility exists of a high Type I error rate with regards to the testing for the clustering of NEC events. The significance levels for individual tests may be extreme if N is large. The danger of rejecting true differences would be the price for limiting the overall error rate. Daily monitoring of rare events would make the number of comparisons so large that prescribed, or adjusted, rejection levels for p-values could become smaller than the extreme data cluster arrangements could possibly yield [164]. We could consider reducing the chance of a false positive by using an alpha level of .01 for the incidence density rate comparisons, instead of the conventional .05. However, any p-value < 0.05 might be enough to raise awareness of the possibility that the number of cases seen within a short time frame has become “too high” and could further warrant action steps to take certain precautions in the NICU. The use of statistical tests without adjusting for multiple testing has been referred to as “signaling” rather than significance testing by Mantel and colleagues; investigators want to be warned of a possible event cluster 5% of the time[207].
Conclusions

An accurate description of the population at risk for necrotizing enterocolitis is clearly of importance for planning interventions and designing comparative studies. A prediction model can help clinicians provide individualized care based on a few clinical characteristics. The multivariable model that was created and validated in the current study for predicting necrotizing enterocolitis is the first of its kind, meeting a need for identifying infants at the highest risk of this devastating disease. This model can be utilized in the clinical setting to help guide clinicians with selecting infants who may benefit most from intervention/preventive trials. Although the current predictions due not have high positive predictive value, they have extremely high negative predictive values, allowing to select out those infants who most likely will not develop NEC. As research continues to make headway with regards to new risk factors for necrotizing enterocolitis, particularly genetic and early laboratory findings, the prediction models presented as a result of this work should be modified to maintain applicability and improve upon their predictive abilities. Operationalizing this prediction model should be strongly considered for continuing validation.

Fundamentally, there is no standard definition of a NEC cluster that might allow an investigator to determine whether a specific instance of several cases closely related in time represents a significant grouping. As a result, clusters (or as they are often called, “outbreaks”) of NEC tend to be defined subjectively; the detection and ascertainment of a cluster lie with observations. With the statistical methodology that has been introduced for the detection of temporal clustering of necrotizing enterocolitis, we now have the ability to monitor the clustering of NEC prospectively. Because the methods are sensitive to the most recent days, they are more powerful than the traditional scan tests for monitoring for purposes of immediate intervention, such as cohorting infants. As research on the clustering phenomena continues, more information on the potential role of environmental factors will come to light, allowing for the implementation of targeted appropriate intervention strategies. Since this is the first study to
look at the occurrence of NEC temporal clustering statistically, we now have a technique for defining clusters and subsequent monitoring. Prospective implementation of this prediction model is warranted.

This study on predicting necrotizing enterocolitis among very low birth weight infants addresses a fundamental need for epidemiologic research in this field. The data is available for generating and testing hypotheses, the resources are available for subsequent study designs, and clinicians and epidemiologists can create strong and productive relationships, which were the building blocks that lead to this study’s accomplishments.
### Tables and Figures

#### Table 1. Bell’s Staging for Necrotizing Enterocolitis (NEC) [8, 9]

<table>
<thead>
<tr>
<th>Classification</th>
<th>Stage 1</th>
<th>Stage II</th>
<th>Stage III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected NEC</td>
<td></td>
<td></td>
<td>Advanced NEC</td>
</tr>
<tr>
<td>Systemic signs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Temperature instability</td>
<td></td>
<td>Stage I signs <em>plus:</em></td>
<td></td>
</tr>
<tr>
<td>• Apnea</td>
<td></td>
<td>• Mild metabolic acidosis</td>
<td></td>
</tr>
<tr>
<td>• Bradycardia</td>
<td></td>
<td>• Mild thrombocytopenia</td>
<td></td>
</tr>
<tr>
<td>• Cyanosis</td>
<td></td>
<td>• Poor perfusion</td>
<td></td>
</tr>
<tr>
<td>• Lethargy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Glucose instability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal signs</td>
<td>Aspirates</td>
<td>Stage I signs <em>plus:</em></td>
<td></td>
</tr>
<tr>
<td>• Mild abdominal distension</td>
<td></td>
<td>• Absent bowel sounds</td>
<td></td>
</tr>
<tr>
<td>• Vomiting</td>
<td></td>
<td>• Abdominal tenderness</td>
<td></td>
</tr>
<tr>
<td>• Positive occult or blood in stools</td>
<td></td>
<td>• Abdominal cellulites or Mass in right lower quadrant</td>
<td></td>
</tr>
<tr>
<td>Radiological signs</td>
<td>Normal</td>
<td>Stage I signs <em>plus:</em></td>
<td></td>
</tr>
<tr>
<td>• Mild intestinal dilatation</td>
<td></td>
<td>• Persistent occult or gross gastrointestinal bleeding</td>
<td></td>
</tr>
<tr>
<td>• Mild ileus</td>
<td></td>
<td>• Intestinal dilatation</td>
<td></td>
</tr>
</tbody>
</table>

Exclude other disorders via bacterial cultures, electrolyte analysis, maternal drug history, coagulation studies, and contrast studies.

Stage I and II signs *plus:* |
- Signs of shock
- Rapid deterioration of vital signs
- Mixed acidosis
- Respiratory compromise
- Hypotension
- Neutropenia
- Peritonitis
- Marked abdominal tenderness
- Marked distension

Stage I and II signs *plus:* |
- Definite ascites
- Pneumoperitoneum
Figure 1: Pathogenesis of NEC consisting of 3 major factors*.

*Adapted from Noerr, 2003. Adv Neonatal Care 3(3): 107-120. [27]
Table 2: Proposed Risk Factors for NEC in Preterm Infants
(In alphabetical order)

<table>
<thead>
<tr>
<th>Prenatal Factors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenatal steroid use</td>
<td>[42]</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>[40]</td>
</tr>
<tr>
<td>Delivery method</td>
<td>[40, 42]</td>
</tr>
<tr>
<td>Hypertensive disease of pregnancy</td>
<td>[40, 43]</td>
</tr>
<tr>
<td>Maternal age</td>
<td>[40, 43]</td>
</tr>
<tr>
<td>Prenatal care</td>
<td>[40]</td>
</tr>
<tr>
<td>PROM</td>
<td>[40]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perinatal/neonatal Factors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth asphyxia</td>
<td>[208]</td>
</tr>
<tr>
<td>Birth weight</td>
<td>[38, 40-43, 45]</td>
</tr>
<tr>
<td>Caffeine 1st 10 days of life</td>
<td>[42]</td>
</tr>
<tr>
<td>Enteral feeding practices</td>
<td>[66]</td>
</tr>
<tr>
<td>Erythropoietin</td>
<td>[42]</td>
</tr>
<tr>
<td>Gender</td>
<td>[40, 43]</td>
</tr>
<tr>
<td>Gestational age</td>
<td>[38, 43, 66]</td>
</tr>
<tr>
<td>Hypothermia</td>
<td>[39]</td>
</tr>
<tr>
<td>Indomethacin exposure</td>
<td>[42]</td>
</tr>
<tr>
<td>Low Apgar score</td>
<td>[38, 40, 42, 208]</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
<td>[38, 42]</td>
</tr>
<tr>
<td>PDA</td>
<td>[38, 43]</td>
</tr>
<tr>
<td>Race</td>
<td>[40, 41]</td>
</tr>
<tr>
<td>RDS / HMD</td>
<td>[37-39]</td>
</tr>
<tr>
<td>SGA / IUGR</td>
<td>[43, 45]</td>
</tr>
<tr>
<td>Perinatal/neonatal Factors (cont.)</td>
<td>References</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>[42, 43]</td>
</tr>
<tr>
<td>Treatment for hypotension</td>
<td>[38, 42]</td>
</tr>
<tr>
<td>Umbilical artery catheter used</td>
<td>[38, 42, 66, 208]</td>
</tr>
<tr>
<td>Studies that found no significant risk factors for NEC</td>
<td>[44, 55]</td>
</tr>
</tbody>
</table>

Abbreviations: PROM, prolonged rupture of membranes; PDA, patent ductus arteriosus; SGA, small for gestational age; IUGR, intrauterine growth retardation; RDS, respiratory distress syndrome; HMD, hyaline membrane disease.

*Risk factors reported from studies conducted to specifically to investigate potential risk factors for NEC among infants. Does not include risk factors previously reported from clinical or observational studies with NEC as one potential outcome.*

*Enteral feeding practices vary among studies. This term is used as an umbrella term for type of feeding (formula vs. human milk), advancement of feeds, age of feeds.*
Table 3: Published NEC outbreaks†

<table>
<thead>
<tr>
<th>Year</th>
<th>Ref</th>
<th>Duration</th>
<th>Season</th>
<th>Cases</th>
<th>Birth weight (g) [GA]a</th>
<th>Nursery settingb</th>
<th>Typical NEC Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>[83]</td>
<td>10 w</td>
<td>Spring</td>
<td>11</td>
<td>1050 – 1900 g</td>
<td>Premature wards</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63 gastroenteritis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>[185]</td>
<td>3 w</td>
<td>NR</td>
<td>20</td>
<td>16 reported</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>“premature”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>[73]e</td>
<td>24 d</td>
<td>NR</td>
<td>6</td>
<td>1590 g</td>
<td>NICU</td>
<td>NR</td>
</tr>
<tr>
<td>1975</td>
<td>[73]e</td>
<td>17 d</td>
<td>NR</td>
<td>12</td>
<td>1590 g</td>
<td>NICU</td>
<td>NR</td>
</tr>
<tr>
<td>1975</td>
<td>[84]e</td>
<td>8 w</td>
<td>Winter</td>
<td>14</td>
<td>2000 g</td>
<td>Infant ICU</td>
<td>11 cases/year</td>
</tr>
<tr>
<td>1976</td>
<td>[81]</td>
<td>8 w</td>
<td>Summer</td>
<td>8</td>
<td>&lt; 2500 g</td>
<td>ICU &amp; NICU</td>
<td>1.2%</td>
</tr>
<tr>
<td>1977</td>
<td>[94]</td>
<td>10 m</td>
<td>Summer/Autumn</td>
<td>42</td>
<td>1497 (528-3700)</td>
<td>NICUc</td>
<td>0.3%</td>
</tr>
<tr>
<td>1979</td>
<td>[85]</td>
<td>8 w</td>
<td>Autumn</td>
<td>10</td>
<td>≥2000 g</td>
<td>NICU</td>
<td>~6 cases/year</td>
</tr>
<tr>
<td>1979</td>
<td>[85]</td>
<td>7 w</td>
<td>Winter</td>
<td>15</td>
<td>≥2000 g</td>
<td>NICU</td>
<td>~11 cases/year</td>
</tr>
<tr>
<td>1979</td>
<td>[85]</td>
<td>20 w</td>
<td>Winter/Spring</td>
<td>15 confirmed</td>
<td>≥2000 g</td>
<td>NICU</td>
<td>0</td>
</tr>
<tr>
<td>1980</td>
<td>[79]</td>
<td>8 w</td>
<td>Winter</td>
<td>12i</td>
<td>1667 g (all premature)</td>
<td>NICU</td>
<td>0 cases / 2 months</td>
</tr>
<tr>
<td>Year</td>
<td>Ref</td>
<td>Duration</td>
<td>Season</td>
<td>Cases</td>
<td>Birth weight (g)</td>
<td>Nursery setting</td>
<td>Typical NEC Rate</td>
</tr>
<tr>
<td>------</td>
<td>-----</td>
<td>----------</td>
<td>--------</td>
<td>-------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1982</td>
<td>[86]</td>
<td>10 w</td>
<td>Spring</td>
<td>32&lt;sup&gt;d&lt;/sup&gt;</td>
<td>NR</td>
<td>[GA]&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NN</td>
</tr>
<tr>
<td>1983</td>
<td>[186]</td>
<td>4 w</td>
<td>Summer</td>
<td>8</td>
<td>1712 ±600</td>
<td>NR</td>
<td>3 gastroenteritis</td>
</tr>
<tr>
<td>1983</td>
<td>[192]</td>
<td>2 w</td>
<td>Autumn</td>
<td>7 confirmed</td>
<td>1800 ± 101.4</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>1984</td>
<td>[87]</td>
<td>8 w</td>
<td>Winter</td>
<td>9 confirmed</td>
<td>1112 ±99.7</td>
<td>NICU</td>
<td>NR</td>
</tr>
<tr>
<td>1984</td>
<td>[187]</td>
<td>4 w</td>
<td>Spring</td>
<td>20</td>
<td>1279 ± 448</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>1985</td>
<td>[88]</td>
<td>8 w</td>
<td>Summer</td>
<td>7</td>
<td>NR</td>
<td>Newborn nursery</td>
<td>4 gastroenteritis</td>
</tr>
<tr>
<td>1987</td>
<td>[89]</td>
<td>10 w</td>
<td>Spring</td>
<td>12 confirmed</td>
<td>1360 g</td>
<td>NICU</td>
<td>.31 cases /1000 patients</td>
</tr>
<tr>
<td>1988</td>
<td>[74]</td>
<td>12 w</td>
<td>NR</td>
<td>7</td>
<td>NR</td>
<td>NICU</td>
<td>NR</td>
</tr>
<tr>
<td>1999</td>
<td>[90]</td>
<td>8 w</td>
<td>Spring</td>
<td>6</td>
<td>1326 ±262.8</td>
<td>PICU</td>
<td>NR</td>
</tr>
<tr>
<td>Year</td>
<td>Ref</td>
<td>Duration</td>
<td>Season</td>
<td>Cases</td>
<td>Birth weight (g) [GA]a</td>
<td>Nursery settingb</td>
<td>Typical NEC Rate</td>
</tr>
<tr>
<td>------</td>
<td>-----</td>
<td>----------</td>
<td>-----------------</td>
<td>------------</td>
<td>------------------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>1999</td>
<td>[95]</td>
<td>3 m</td>
<td>Winter</td>
<td>19</td>
<td>500 - &gt;2500 g</td>
<td>NICU &amp; ICN</td>
<td>6.8%</td>
</tr>
<tr>
<td>2000</td>
<td>[91]</td>
<td>9 w</td>
<td>Summer/Autumn</td>
<td>4</td>
<td>3000 g (2800-3400)</td>
<td>PICU</td>
<td>0</td>
</tr>
<tr>
<td>2001</td>
<td>[93]</td>
<td>8 w</td>
<td>Summer</td>
<td>12</td>
<td>&lt; 2000 g</td>
<td>NICU</td>
<td>NR</td>
</tr>
<tr>
<td>2002</td>
<td>[92]</td>
<td>8 w</td>
<td>Winter</td>
<td>6 confirmed</td>
<td>&gt;30 weeks</td>
<td>NICU &amp; ICN</td>
<td>1-2 cases/year</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 suspected</td>
</tr>
<tr>
<td>2004</td>
<td>[96]</td>
<td>5 w</td>
<td>Summer</td>
<td>16 confirmed</td>
<td>1325 – 3615 g</td>
<td>2 units</td>
<td>0.1 cases /1000 per year</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 suspected</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12 gastroenteritis</td>
</tr>
</tbody>
</table>

Abbreviations: NR = Not Reported, w=weeks, m=months, d=days, g=grams.

aGestational age (GA) reported if birth weight not available. Birth weight reported as mean ±SD or median (range) if available.

bNICU=Neonatal intensive care unit, ICN = Intermediate care nursery, NN = newborn nursery.

cSame hospital, but 2 different NICU rooms.

dTwo hospitals included in investigation. All NEC cases occurred at one hospital setting.

eReported in abstract form only. f11/12 cases occurred in an 18-day period.
### Table 4: Risk factors reported with the outbreaks†

<table>
<thead>
<tr>
<th>Reference</th>
<th>Risk Factors (predisposing factors)*</th>
<th>Pathogen detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>[83]</td>
<td>Prematurity, neonatal stress</td>
<td><em>Salmonella thompson or johannesburg</em></td>
</tr>
<tr>
<td>[73]</td>
<td>No significant correlation with risk factors.</td>
<td>Enteric forms in stool</td>
</tr>
<tr>
<td>[84]</td>
<td>Low apgar, age at first enteral feed</td>
<td><em>Klebsiella</em></td>
</tr>
<tr>
<td>[81]</td>
<td>Low birth weight, gestational age, low apgar, perinatal complications,</td>
<td><em>E. coli</em> not typable</td>
</tr>
<tr>
<td>[94]</td>
<td>NR</td>
<td>No significant pathogen</td>
</tr>
<tr>
<td>[85]</td>
<td>No significant correlation with risk factors.</td>
<td><em>Klebsiella pneumoniae</em> in 1 of 3 outbreaks.</td>
</tr>
<tr>
<td>[79]</td>
<td>Prematurity, hyaline membrane disease, preinatal shock, umbilical catheter</td>
<td><em>Enterobacter cloacae</em> type 3305573</td>
</tr>
<tr>
<td>[86]</td>
<td>No significant correlation with risk factors</td>
<td>Coronavirus, rotavirus</td>
</tr>
<tr>
<td>[186]</td>
<td>No significant correlation with risk factors.</td>
<td>Rotavirus</td>
</tr>
<tr>
<td>[192]</td>
<td>NR</td>
<td><em>E. coli</em> heat labile toxin</td>
</tr>
<tr>
<td>[87]</td>
<td>Low birth weight, maternal toxemia</td>
<td>No specific pathogen</td>
</tr>
<tr>
<td>[187]</td>
<td>Chronological age</td>
<td>No specific pathogen</td>
</tr>
<tr>
<td>Reference</td>
<td>Risk Factors (predisposing factors)a</td>
<td>Pathogen detected</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>[88]</td>
<td>NR</td>
<td>No specific pathogen</td>
</tr>
<tr>
<td>[89]</td>
<td>Low birth weight, transfusion of packed red blood cells, use of diuretics</td>
<td>No significant pathogen</td>
</tr>
<tr>
<td>[74]</td>
<td>Low birth weight, age at first enteral feed</td>
<td>Rotavirus</td>
</tr>
<tr>
<td>[90]</td>
<td>NR</td>
<td><em>Klebsiella pneumoniae</em></td>
</tr>
<tr>
<td>[95]</td>
<td>Low birth weight</td>
<td>No significant pathogen b</td>
</tr>
<tr>
<td>[91]</td>
<td>Higher mean temperature gradient</td>
<td>No significant pathogen</td>
</tr>
<tr>
<td>[93]</td>
<td>Feeds from powdered milk formula</td>
<td><em>Enterobacter sakazakii</em></td>
</tr>
<tr>
<td>[92]</td>
<td>NR</td>
<td><em>Clostridium neonatale</em></td>
</tr>
<tr>
<td>[96]</td>
<td>Invasive procedures, pathological conditions, age at first feedings with formula</td>
<td>NR</td>
</tr>
</tbody>
</table>

†NR = Not Reported

aRisk factors reported from either case-control study design or case reports.

bAuthors reported pathogens specific to investigation of nosocomial infections, but none directly associated with NEC.
Figure 2: Example of generic classification and regression tree output*

*Adapted from Lemon et al. [128]
Figure 3: Birth weight cut-off for stratified analysis. Cut-off based on the best sensitivity-specificity trade off.

*Sensitivity for this cut off is 71% and specificity is 51%.
Figure 4: Proportion of VLBW infants who are small for gestational age [120]
Figure 5: Incidence and 95% Confidence Interval of Necrotizing Enterocolitis by Year*

*Dashed horizontal line indicates the mean rate of NEC for all years combined (6.96%)*
Figure 6: Incidence and 95% Confidence Interval of NEC by birth weight categories*

*Incidence of NEC decreasing with increasing birth weight (p<.0001 test for trend).
Figure 7: Scatter plot and Regression Line of Age of NEC Onset and Birth Weight

Regression Line ($\beta = -0.013, p<.0001$) with 95% Confidence Intervals.
Figure 8: Median and Interquartile Ranges of NEC by Birth Weight
Figure 9: Mortality rates over time for infants with and without NEC
Table 5: Comparison of Maternal/Pregnancy Characteristics between Derivation and Validation Samples

<table>
<thead>
<tr>
<th></th>
<th>Derivation N=3435</th>
<th>Validation N=381</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>26.1 (6.6)</td>
<td>26.4 (6.7)</td>
<td>0.41</td>
</tr>
<tr>
<td>Maternal age &lt; 21 y</td>
<td>852 (25)</td>
<td>92 (24)</td>
<td>0.78</td>
</tr>
<tr>
<td>Marital status – single (65 missing)</td>
<td>1659 (48)</td>
<td>180 (48)</td>
<td>0.63</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>3234 (94)</td>
<td>358 (94)</td>
<td>0.93</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>2131 (62)</td>
<td>231 (61)</td>
<td>0.56</td>
</tr>
<tr>
<td>Diabetes/insulin dependentb</td>
<td>96 (4)</td>
<td>14 (6)</td>
<td>0.43</td>
</tr>
<tr>
<td>Hypertension</td>
<td>877 (26)</td>
<td>104 (27)</td>
<td>0.48</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>674 (20)</td>
<td>78 (20)</td>
<td>0.71</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>559 (16)</td>
<td>52 (14)</td>
<td>0.18</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>1847 (54)</td>
<td>207 (54)</td>
<td>0.80</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>2441 (71)</td>
<td>275 (72)</td>
<td>0.47</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>1292 (38)</td>
<td>150 (39)</td>
<td>0.50</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>2168 (63)</td>
<td>233 (61)</td>
<td>0.45</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>976 (28)</td>
<td>115 (30)</td>
<td>0.47</td>
</tr>
</tbody>
</table>

aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

bMissing data on 1413 subjects (36% of data).
Table 6: Comparison of Clinical Characteristics between Derivation and Validation Samples

<table>
<thead>
<tr>
<th></th>
<th>Derivation N=3435</th>
<th>Validation N=381</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender – Male</td>
<td>1694 (49)</td>
<td>177 (46)</td>
<td>0.29</td>
</tr>
<tr>
<td>Race – African American</td>
<td>1140 (33)</td>
<td>115 (30)</td>
<td>0.24</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>1058 (279)</td>
<td>1055 (286)</td>
<td>0.86</td>
</tr>
<tr>
<td>Birth weight &lt; 1000 grams</td>
<td>1444 (42)</td>
<td>151 (40)</td>
<td>0.37</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>28.2 (2.7)</td>
<td>28.2 (2.7)</td>
<td>0.99</td>
</tr>
<tr>
<td>SGA</td>
<td>772 (22)</td>
<td>89 (23)</td>
<td>0.70</td>
</tr>
<tr>
<td>Onset of NEC</td>
<td>231 (7)</td>
<td>32 (8)</td>
<td>0.22</td>
</tr>
<tr>
<td>Age at onset of NEC in days</td>
<td>18 (14)</td>
<td>22 (8)</td>
<td>0.13</td>
</tr>
<tr>
<td>5-minute Apgar Score</td>
<td>7.1 (2.1)</td>
<td>7.0 (2.1)</td>
<td>0.45</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>1017 (30)</td>
<td>118 (31)</td>
<td>0.58</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of life</td>
<td>1211 (35)</td>
<td>123 (32)</td>
<td>0.25</td>
</tr>
<tr>
<td>Number of days on ventilator first week of life</td>
<td>2.3 (3.1)</td>
<td>2.3 (3.0)</td>
<td>0.66</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>1437 (42)</td>
<td>157 (41)</td>
<td>0.81</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>966 (28)</td>
<td>109 (29)</td>
<td>0.84</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus (PDA)</td>
<td>1182 (34)</td>
<td>140 (37)</td>
<td>0.36</td>
</tr>
<tr>
<td>PDA Treatmentc</td>
<td>938 (27)</td>
<td>108 (28)</td>
<td>0.67</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>144 (4)</td>
<td>20 (5)</td>
<td>0.33</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>1810 (53)</td>
<td>196 (51)</td>
<td>0.64</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>1307 (38)</td>
<td>148 (39)</td>
<td>0.79</td>
</tr>
<tr>
<td>Required O2 at 6 hrs thru 24 hrs of age&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2040 (64)</td>
<td>233 (67)</td>
<td>0.42</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of age&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2824 (89)</td>
<td>313 (89)</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>Derivation N=3435</td>
<td>Validation N=381</td>
<td>P-value</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrs(^d)</td>
<td>2178 (69)</td>
<td>239 (68)</td>
<td>0.86</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrs(^d)</td>
<td>2332 (74)</td>
<td>254 (73)</td>
<td>0.68</td>
</tr>
<tr>
<td>High FIO2 (&gt;0.9)(^e)</td>
<td>82 (4)</td>
<td>4 (2)</td>
<td>0.086</td>
</tr>
<tr>
<td>Low temperature (&lt;35(^\circ))(^e)</td>
<td>229 (23)</td>
<td>30 (27)</td>
<td>0.35</td>
</tr>
<tr>
<td>Received 1(^st) enteral feed ≤ 3 days of age</td>
<td>2084 (61)</td>
<td>238 (62)</td>
<td>0.50</td>
</tr>
<tr>
<td>Achieved full feeds during NICU stay</td>
<td>3098 (90)</td>
<td>344 (90)</td>
<td>0.97</td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds(^f)</td>
<td>14.6 (11.3)</td>
<td>14.9 (12.2)</td>
<td>0.70</td>
</tr>
<tr>
<td>Full enteral feeds by 5 days of age</td>
<td>336 (10)</td>
<td>31 (8)</td>
<td>0.30</td>
</tr>
<tr>
<td>Full enteral feeds by 7 days of age</td>
<td>800 (23)</td>
<td>91 (24)</td>
<td>0.79</td>
</tr>
<tr>
<td>Full enteral feeds by 14 days of age</td>
<td>1980 (58)</td>
<td>234 (61)</td>
<td>0.16</td>
</tr>
</tbody>
</table>

\(^a\)Categorical variables reported as n (%) and continuous variables reported as mean (SD).

\(^b\)For years <1998, on conventional ventilator for first 10 days of life; for ≥1998, on conventional ventilator for first 7 days of life.

\(^c\)PDA treatment with indomethacin and/or surgery.

\(^d\)Missing data on 298 subjects (8% of data).

\(^e\)Missing data FIO2 data on 1515 subjects (40% of data) and missing temperature data on 2713 subjects (71% of data).

\(^f\)Among infants who achieved full enteral feeding during NICU stay.
Table 7: Maternal/Pregnancy Characteristics of Derivation Sample

<table>
<thead>
<tr>
<th></th>
<th>Derivation (n=3435)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NEC N=231</td>
<td>No NEC N=3204</td>
<td>OR or mean difference</td>
<td>95% CI</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td>Maternal age</td>
<td>25.8 (6.7)</td>
<td>26.1 (6.6)</td>
<td>0.37</td>
<td>-0.46, 1.2</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Maternal age &lt; 21 yrs</td>
<td>72 (27)</td>
<td>872 (25)</td>
<td>1.2</td>
<td>0.87, 1.5</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Marital status – single</td>
<td>123 (54)</td>
<td>1536 (49)</td>
<td>1.2</td>
<td>0.94, 1.6</td>
<td>0.136</td>
<td></td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>219 (95)</td>
<td>3015 (94)</td>
<td>1.3</td>
<td>.71, 2.5</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>135 (59)</td>
<td>1996 (62)</td>
<td>0.86</td>
<td>0.65, 1.1</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Diabetes/insulin dependentb</td>
<td>8 (5)</td>
<td>88 (4)</td>
<td>1.2</td>
<td>0.59, 2.6</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>48 (21)</td>
<td>829 (26)</td>
<td>0.75</td>
<td>0.54, 1.04</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>44 (19)</td>
<td>630 (20)</td>
<td>0.96</td>
<td>0.68, 1.3</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>47 (20)</td>
<td>512 (16)</td>
<td>1.3</td>
<td>0.96, 1.9</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>139 (60)</td>
<td>1708 (53)</td>
<td>1.3</td>
<td>1.01, 1.7</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>172 (74)</td>
<td>2269 (71)</td>
<td>1.2</td>
<td>0.88, 1.6</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>81 (35)</td>
<td>1211 (38)</td>
<td>0.89</td>
<td>0.67, 1.2</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>159 (69)</td>
<td>2009 (63)</td>
<td>1.3</td>
<td>0.99, 1.8</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Multiple birth</td>
<td>63 (27)</td>
<td>913 (28.5)</td>
<td>0.94</td>
<td>0.70, 1.3</td>
<td>0.069</td>
<td></td>
</tr>
</tbody>
</table>

aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

bMissing values on 1399 subjects (37% of data).
Table 8: Clinical characteristics of VLBW infants for Derivation Samplea

<table>
<thead>
<tr>
<th>Predictor</th>
<th>NEC</th>
<th>No NEC</th>
<th>OR or mean difference</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender – Male</td>
<td>122 (53)</td>
<td>1572 (49)</td>
<td>1.2</td>
<td>.89, 1.5</td>
<td>0.27</td>
</tr>
<tr>
<td>Race – African American</td>
<td>96 (42)</td>
<td>1044 (33)</td>
<td>1.5</td>
<td>1.1, 1.9</td>
<td>0.005</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>26.9 (2.2)</td>
<td>28.3 (2.7)</td>
<td>1.4</td>
<td>1.0, 1.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>935 (266)</td>
<td>1067 (278)</td>
<td>132</td>
<td>98, 167</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>SGA</td>
<td>38 (16)</td>
<td>734 (23)</td>
<td>0.66</td>
<td>.46, .95</td>
<td>0.02</td>
</tr>
<tr>
<td>Age at onset of NEC in days</td>
<td>17.9 (13.3)</td>
<td>---------</td>
<td>-----------------------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>5-min Apgar Score</td>
<td>6.8 (2.1)</td>
<td>7.1 (2.1)</td>
<td>0.30</td>
<td>0.04, .56</td>
<td>0.023</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>82 (36)</td>
<td>935 (29)</td>
<td>1.3</td>
<td>1.0, 1.8</td>
<td>0.042</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of life</td>
<td>112 (48)</td>
<td>1000 (34)</td>
<td>1.8</td>
<td>1.3, 2.4</td>
<td>0.0001</td>
</tr>
<tr>
<td>Number of days on ventilator first week of life</td>
<td>3.0 (3.2)</td>
<td>2.3 (3.1)</td>
<td>-0.70</td>
<td>-1.1,-.31</td>
<td>0.0004</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>128 (55)</td>
<td>1309 (41)</td>
<td>1.8</td>
<td>1.4, 2.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>83 (36)</td>
<td>883 (28)</td>
<td>1.4</td>
<td>1.1, 1.9</td>
<td>0.018</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus</td>
<td>112 (48)</td>
<td>1070 (33)</td>
<td>1.9</td>
<td>1.4, 2.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Treatmentc</td>
<td>85 (37)</td>
<td>853 (27)</td>
<td>1.6</td>
<td>1.2, 2.1</td>
<td>0.0008</td>
</tr>
<tr>
<td>Event</td>
<td>NEC N=231</td>
<td>No NEC N=3204</td>
<td>OR or mean difference</td>
<td>95% CI</td>
<td>p-value</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------</td>
<td>---------------</td>
<td>-----------------------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>26 (11)</td>
<td>118 (4)</td>
<td>3.3</td>
<td>2.1, 5.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>135 (58)</td>
<td>1675 (52)</td>
<td>1.3</td>
<td>0.98, 1.7</td>
<td>0.07</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>123 (54)</td>
<td>1184 (37)</td>
<td>1.9</td>
<td>1.5, 2.6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Required O₂ at 6 hrs thru 24 hrs of age</td>
<td>150 (69)</td>
<td>1890 (64)</td>
<td>1.3</td>
<td>0.93, 1.7</td>
<td>0.132</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of age</td>
<td>207 (95)</td>
<td>2617 (89)</td>
<td>2.6</td>
<td>1.4, 5.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Need for resp. support 24 hrs</td>
<td>165 (76)</td>
<td>2013 (68)</td>
<td>1.5</td>
<td>1.1, 2.0</td>
<td>0.016</td>
</tr>
<tr>
<td>Abnormal chest x-ray 24 hrs</td>
<td>178 (82)</td>
<td>2154 (73)</td>
<td>1.7</td>
<td>1.2, 2.4</td>
<td>0.004</td>
</tr>
<tr>
<td>High FIO₂ (&gt;0.9)</td>
<td>12 (9%)</td>
<td>70 (4%)</td>
<td>2.5</td>
<td>1.3, 4.7</td>
<td>0.004</td>
</tr>
<tr>
<td>Low temperature at admission (&lt;35°)</td>
<td>20 (35)</td>
<td>209 (22%)</td>
<td>1.9</td>
<td>1.1, 3.3</td>
<td>0.027</td>
</tr>
<tr>
<td>Received 1st enteral feed ≤ 3 d of age</td>
<td>126 (55)</td>
<td>1958 (61)</td>
<td>0.76</td>
<td>0.58, 0.99</td>
<td>0.05</td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds</td>
<td>22.4 (17)</td>
<td>14.1 (11)</td>
<td>-8.3</td>
<td>-9.8, -6.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>NEC N=231</td>
<td>No NEC N=3204</td>
<td>OR or mean difference</td>
<td>95% CI</td>
<td>p-value</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------</td>
<td>---------------</td>
<td>-----------------------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>Full enteral feeds by 5 d of age</td>
<td>4 (2%)</td>
<td>332 (10%)</td>
<td>0.15</td>
<td>0.06, 0.41</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 7d of age</td>
<td>22 (10)</td>
<td>778 (24)</td>
<td>0.33</td>
<td>0.21, 0.51</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 14 d of age</td>
<td>77 (33)</td>
<td>1903 (59)</td>
<td>0.34</td>
<td>0.26, 0.45</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

*bFor years <1998, on conventional ventilator for first 10 days of life; for ≥1998, on conventional ventilator for first 7 days of life.

*cPDA (patent ductus arteriosus) treatment with indomethacin and/or surgery.

*dMissing values for 298 subjects (8% of data).

*eMissing data FIO2 data on 1515 subjects (40% of data) and missing temperature data on 2713 subjects (71% of data).

*fAmong infants who achieved full enteral feeding during NICU stay.
Table 9: Logistic regression results from derivation sample (n=3435)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β coefficient</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.76</td>
<td>0.33</td>
<td>&lt;.0001</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.0013</td>
<td>0.0003</td>
<td>&lt;.0001</td>
<td>0.88a</td>
<td>0.84, 0.93</td>
</tr>
<tr>
<td>African American</td>
<td>0.33</td>
<td>0.14</td>
<td>0.018</td>
<td>1.40</td>
<td>1.06, 1.84</td>
</tr>
<tr>
<td>Non Hispanic White REF</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>-1.39</td>
<td>0.52</td>
<td>0.007</td>
<td>0.25</td>
<td>0.09, 0.69</td>
</tr>
<tr>
<td>PDA(^b) No treatment</td>
<td>0.56</td>
<td>0.23</td>
<td>0.015</td>
<td>1.75</td>
<td>1.11, 2.75</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>-0.008</td>
<td>0.18</td>
<td>0.96</td>
<td>0.99</td>
<td>0.70, 1.40</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.01</td>
<td>0.25</td>
<td>&lt;.001</td>
<td>2.76</td>
<td>1.70, 4.47</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>0.31</td>
<td>0.15</td>
<td>0.036</td>
<td>1.37</td>
<td>1.02, 1.83</td>
</tr>
</tbody>
</table>

\(AUC = 0.682 \ (95\% \ CI \ 0.648 – 0.715)\)

Goodness-of-fit statistics: Hosmer - Lemeshow p=0.56

Deviance p=1.0 (Value/DF = 0.55); Pearson p= 0.20 (Value/DF = 1.03)

\(^a\)With every 100 gram increase in birth weight, likelihood of NEC decreases.

\(^b\)PDA = patent ductus arteriosus.
Table 10: Logistic regression results Derivation Sample with interaction term (n=3435)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.58</td>
<td>0.49</td>
<td>&lt;.0001</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Birth weight(^a)</td>
<td>-0.0004</td>
<td>0.0005</td>
<td>0.39</td>
<td>----</td>
<td>--------</td>
</tr>
<tr>
<td>African American</td>
<td>0.34</td>
<td>0.14</td>
<td>0.017</td>
<td>1.40</td>
<td>1.06, 1.85</td>
</tr>
<tr>
<td>Non Hispanic White</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding</td>
<td>-1.39</td>
<td>0.52</td>
<td>0.007</td>
<td>0.25</td>
<td>0.09, 0.69</td>
</tr>
<tr>
<td>by 5 d of age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDA(^b) No treatment</td>
<td>0.55</td>
<td>0.23</td>
<td>0.017</td>
<td>1.73</td>
<td>1.10, 2.72</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>-0.008</td>
<td>0.18</td>
<td>0.96</td>
<td>0.99</td>
<td>0.70, 1.40</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.04</td>
<td>0.25</td>
<td>&lt;.0001</td>
<td>2.82</td>
<td>1.74, 4.57</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>1.53</td>
<td>0.54</td>
<td>0.005</td>
<td>----</td>
<td>--------</td>
</tr>
<tr>
<td>Birth weight x parity</td>
<td>-0.0013</td>
<td>0.0005</td>
<td>0.018</td>
<td>----</td>
<td>--------</td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing weight</td>
<td></td>
<td></td>
<td></td>
<td>1.18</td>
<td>1.11, 1.26</td>
</tr>
<tr>
<td>Parity = 1, decreasing weight</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
<td>1.04</td>
<td>0.95, 1.14</td>
</tr>
</tbody>
</table>

\(^a\)Interpretation of birth weight occurs when interpreting the interaction between birth weight and parity. Among infants who are not the first born (parity > 1), for every 100 gram decrease in birth weight, the likelihood of NEC increases by 1.18.

\(^b\)PDA = Patent Ductus Arteriosus.

\(AUC = 0.685 \ (95\% \ CI \ 0.652 – 0.719 )\)

Goodness-of-fit statistics: H-L p=0.27

Deviance p=1.0 (Value/DF = 0.55) Pearson p= 0.83 (Value/DF = 0.97)
Figure 10: Calibration plot for derivation model

Hosmer-Lemeshow \( p = 0.27 \)
Table 11: Logistic regression results Validation Sample (n=381)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.60</td>
<td>1.32</td>
<td>0.05</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Birth weight&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.0009</td>
<td>0.001</td>
<td>0.45</td>
<td>----</td>
<td>--------</td>
</tr>
<tr>
<td>African American</td>
<td>1.0005</td>
<td>0.40</td>
<td>0.01</td>
<td>2.72</td>
<td>1.25, 5.91</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>0.92</td>
<td>0.72</td>
<td>0.20</td>
<td>2.51</td>
<td>0.62, 10.21</td>
</tr>
<tr>
<td>PDA&lt;sup&gt;b&lt;/sup&gt; No treatment</td>
<td>1.72</td>
<td>0.58</td>
<td>0.003</td>
<td>5.6</td>
<td>1.80, 17.44</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>1.07</td>
<td>0.52</td>
<td>0.04</td>
<td>2.91</td>
<td>1.05, 8.01</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.09</td>
<td>0.76</td>
<td>0.15</td>
<td>2.97</td>
<td>0.67, 13.19</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>0.38</td>
<td>1.41</td>
<td>0.79</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Birth weight x parity</td>
<td>-0.0003</td>
<td>0.001</td>
<td>0.83</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing weight</td>
<td>----------------</td>
<td>0.17</td>
<td>1.13</td>
<td>0.95, 1.35</td>
<td></td>
</tr>
<tr>
<td>Parity = 1, decreasing weight</td>
<td>----------------</td>
<td>0.45</td>
<td>1.10</td>
<td>0.86, 1.40</td>
<td></td>
</tr>
</tbody>
</table>

AUC = 0.734 (95% CI 0.642 – 0.830 )

Goodness-of-fit statistics: H-L p=0.28

Deviance p=1.0 (Value/DF = 0.54); Pearson p= 0.25 (Value/DF = 1.06)

<sup>a</sup>Interpretation of birth weight occurs when interpreting the interaction between birth weight and parity.

<sup>b</sup>PDA = Patent Ductus Arteriosus.
Figure 11: Area under the ROC curve for Derivation and Validation Models

ROC = receiver operator characteristic; AUC = area under the curve
Table 12: Associations between clinical characteristics and the presence of NEC in the derivation sample and the validation sample

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Derivation sample</th>
<th>Validation sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=3435)</td>
<td>(n=381)</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Birth weight (per 100 gram increase)(^b)</td>
<td>-------</td>
<td>----------------</td>
</tr>
<tr>
<td>African American</td>
<td>1.40</td>
<td>1.06, 1.85</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>0.25</td>
<td>0.09, 0.69</td>
</tr>
<tr>
<td>PDA(^c) No treatment</td>
<td>1.73</td>
<td>1.10, 2.72</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>0.99</td>
<td>0.70, 1.40</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>2.82</td>
<td>1.74, 4.57</td>
</tr>
<tr>
<td>No PDA reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity x birth weight(^b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing weight</td>
<td>1.18</td>
<td>1.11, 1.26</td>
</tr>
<tr>
<td>Parity = 1, decreasing weight</td>
<td>1.04</td>
<td>0.95, 1.14</td>
</tr>
<tr>
<td>AUC</td>
<td>0.685</td>
<td>.652, .719</td>
</tr>
</tbody>
</table>

\(^a\)Data presented as multivariable odds ratio (95% confidence interval).

\(^b\)Results of birth weight are interpreted in the interaction term between infant parity and infant birth weight.

\(^c\)PDA = Patent Ductus Arteriosus.
Figure 12: Calibration plots for both Derivation and Validation Models

H-L = Hosmer-Lemeshow
Table 13: Maternal/Pregnancy Characteristics; All infants\(^a\) (n=3816)

<table>
<thead>
<tr>
<th></th>
<th>NEC</th>
<th>No NEC</th>
<th>OR or mean difference (SE)</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>25.8 (6.7)</td>
<td>26.1 (6.6)</td>
<td>0.37 (.42)</td>
<td>(-0.46, 1.2)</td>
<td>.279</td>
</tr>
<tr>
<td>Maternal age &lt; 21 yrs</td>
<td>72 (27)</td>
<td>872 (25)</td>
<td>1.2</td>
<td>0.87, 1.5</td>
<td>.38</td>
</tr>
<tr>
<td>Marital status – single</td>
<td>140 (54)</td>
<td>1699 (49)</td>
<td>1.2</td>
<td>0.95, 1.6</td>
<td>.11</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>250 (95)</td>
<td>3342 (94)</td>
<td>1.23</td>
<td>0.69, 2.18</td>
<td>.484</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>154 (59)</td>
<td>2208 (62)</td>
<td>0.87</td>
<td>0.67, 1.1</td>
<td>.26</td>
</tr>
<tr>
<td>Diabetes/insulin dependent(^b)</td>
<td>10 (6)</td>
<td>100 (4)</td>
<td>1.3</td>
<td>0.67, 2.6</td>
<td>.42</td>
</tr>
<tr>
<td>Hypertension</td>
<td>54 (21)</td>
<td>927 (26)</td>
<td>0.73</td>
<td>0.53, 0.99</td>
<td>.043</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>54 (21)</td>
<td>698 (20)</td>
<td>1.1</td>
<td>0.77, 1.4</td>
<td>.74</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>55 (21)</td>
<td>556 (16)</td>
<td>1.4</td>
<td>1.0, 1.9</td>
<td>.025</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>158 (60)</td>
<td>1896 (53)</td>
<td>1.3</td>
<td>1.0, 1.7</td>
<td>.036</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>200 (76)</td>
<td>2516 (71)</td>
<td>1.3</td>
<td>0.97, 1.7</td>
<td>.077</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>95 (36)</td>
<td>1347 (38)</td>
<td>0.93</td>
<td>0.71, 1.2</td>
<td>.56</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>179 (68)</td>
<td>2222 (63)</td>
<td>1.3</td>
<td>0.98, 1.7</td>
<td>.074</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>75 (29)</td>
<td>1016 (29)</td>
<td>1.0</td>
<td>0.75, 1.31</td>
<td>.98</td>
</tr>
</tbody>
</table>

\(^a\)Categorical variables reported as n (%) and continuous variables reported as mean (SD).

\(^b\)Missing values on 1399 subjects (37% of the data)
<table>
<thead>
<tr>
<th></th>
<th>NEC N=263</th>
<th>No NEC N=3553</th>
<th>OR or mean difference (SE)</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender – Male</td>
<td>139 (52)</td>
<td>1732 (49)</td>
<td>1.2</td>
<td>0.92, 1.5</td>
<td>0.20</td>
</tr>
<tr>
<td>Race – African American</td>
<td>112 (43)</td>
<td>1143 (32)</td>
<td>1.6</td>
<td>1.2, 2.0</td>
<td>0.0005</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>26.9 (2.2)</td>
<td>28.3 (2.7)</td>
<td>1.4 (.17)</td>
<td>1.01, 1.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>934 (266)</td>
<td>1067 (278)</td>
<td>132 (18)</td>
<td>98, 167</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>SGA</td>
<td>43 (16)</td>
<td>818 (23)</td>
<td>0.65</td>
<td>0.47, 0.91</td>
<td>0.013</td>
</tr>
<tr>
<td>5-minute Apgar Score</td>
<td>6.8 (2.1)</td>
<td>7.1 (2.1)</td>
<td>0.30 (.13)</td>
<td>0.04, 0.56</td>
<td>.02</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>90 (34)</td>
<td>1045 (29)</td>
<td>1.26</td>
<td>0.96, 1.6</td>
<td>0.099</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of life</td>
<td>123 (47)</td>
<td>1211 (34)</td>
<td>1.7</td>
<td>1.3, 2.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Number of days on ventilator first week of lifeb</td>
<td>3.0 (3.2)</td>
<td>2.3 (3.1)</td>
<td>-0.70 (.20)</td>
<td>-1.1, -0.31</td>
<td>0.0004</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>141 (54)</td>
<td>1453 (41)</td>
<td>1.7</td>
<td>1.3, 2.1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>96 (37)</td>
<td>979 (28)</td>
<td>1.5</td>
<td>1.2, 2.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus (PDA)</td>
<td>133 (51)</td>
<td>1189 (33)</td>
<td>2.0</td>
<td>1.6, 2.6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Treatmentc</td>
<td>99 (38)</td>
<td>947 (27)</td>
<td>1.7</td>
<td>1.3, 2.2</td>
<td>0.0001</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>29 (11)</td>
<td>135 (4)</td>
<td>3.1</td>
<td>2.1, 4.8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>148 (56)</td>
<td>1858 (52)</td>
<td>1.2</td>
<td>0.91, 1.5</td>
<td>0.21</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>134 (51)</td>
<td>1321 (37)</td>
<td>1.8</td>
<td>1.4, 2.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Required O2 at 6 hrs thru 24 hrsd</td>
<td>174 (70)</td>
<td>2099 (64)</td>
<td>1.3</td>
<td>0.99, 1.7</td>
<td>0.058</td>
</tr>
<tr>
<td>Showed resp. distress within 24hrsd</td>
<td>237 (96)</td>
<td>2900 (89)</td>
<td>2.7</td>
<td>1.5, 5.1</td>
<td>0.0008</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrsd</td>
<td>187 (75)</td>
<td>2230 (68)</td>
<td>1.4</td>
<td>1.1, 1.9</td>
<td>0.018</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrsd</td>
<td>201 (81)</td>
<td>2385 (73)</td>
<td>1.6</td>
<td>1.1, 2.2</td>
<td>0.005</td>
</tr>
<tr>
<td>High FIO2 (&gt;0.9)e</td>
<td>12 (8)</td>
<td>74 (3)</td>
<td>2.3</td>
<td>1.2, 4.3</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>NEC N=263</td>
<td>No NEC N=3553</td>
<td>OR or mean difference (SE)</td>
<td>95% CI</td>
<td>p-value</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------</td>
<td>---------------</td>
<td>----------------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Low temperature at admission (&lt;35°)²</td>
<td>24 (34)</td>
<td>235 (23)</td>
<td>1.8</td>
<td>1.1, 2.96</td>
<td>0.028</td>
</tr>
<tr>
<td>Received 1st enteral feed ≤ 3 d of age</td>
<td>154 (57)</td>
<td>2224 (61)</td>
<td>.86</td>
<td>.67, 1.10</td>
<td>.235</td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds³</td>
<td>22.6 (16.9)</td>
<td>14.1 (10.7)</td>
<td>8.5 (.78)</td>
<td>(6.9, 10)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 5 d of age</td>
<td>7 (3)</td>
<td>369 (10)</td>
<td>.24</td>
<td>.11, .51</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

²Categorical variables reported as n (%) and continuous variables reported as mean (SD).

³For years <1998, on conventional ventilator for first 10 days of life; for ≥1998, on conventional ventilator for first 7 days of life.

³PDA treatment with indomethacin and/or surgery.

³Missing data on 298 subjects (8% of data).

⁵Missing data FIO2 data on 1515 subjects (40% of data) and missing temperature data on 2713 subjects (71% of data).

⁶Among infants who achieved full enteral feeding during NICU stay.
Table 15: Logistic regression results on total sample\(^a\) (n=3816)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>(\beta)</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.56</td>
<td>0.45</td>
<td>&lt;.0001</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Birth weight(^b)</td>
<td>-0.0005</td>
<td>0.0004</td>
<td>0.28</td>
<td>----</td>
<td>--------</td>
</tr>
<tr>
<td>African American</td>
<td>0.41</td>
<td>0.13</td>
<td>0.002</td>
<td>1.51</td>
<td>1.16, 1.96</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding</td>
<td>-0.91</td>
<td>0.40</td>
<td>0.02</td>
<td>0.40</td>
<td>0.19, 0.88</td>
</tr>
<tr>
<td>by 5 d of age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDA(^c) No treatment</td>
<td>0.69</td>
<td>0.21</td>
<td>0.001</td>
<td>1.99</td>
<td>1.31, 3.01</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>0.10</td>
<td>0.17</td>
<td>0.56</td>
<td>1.10</td>
<td>0.80, 1.52</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.04</td>
<td>0.23</td>
<td>&lt;.0001</td>
<td>2.83</td>
<td>1.79, 4.49</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>1.43</td>
<td>0.50</td>
<td>0.005</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Birth weight x parity</td>
<td>-0.0012</td>
<td>0.0005</td>
<td>0.017</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing weight</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
<td>1.18(^d)</td>
<td>1.11, 1.26</td>
</tr>
<tr>
<td>Parity = 1, decreasing weight</td>
<td></td>
<td></td>
<td>0.28</td>
<td>1.05</td>
<td>0.96, 1.14</td>
</tr>
</tbody>
</table>

\(AUC = 0.689\) (95% CI 0.658 – 0.720)

Goodness-of-fit statistics: Hosmer – Lemeshow p=0.26; Deviance p=1.0 (Value/DF = 0.55); Pearson p= 0.95 (Value/DF = 0.96)

\(^a\)Data presented as multivariable odds ratio (95% confidence interval).

\(^b\)Results of birth weight are interpreted in the interaction term between infant parity and infant birth weight. \(^c\)PDA = patent ductus arteriosus.

\(^d\)Among infants who are not the first born (parity > 1), for every 100 gram decrease in birth weight, the likelihood of NEC increases by 1.18.
Table 16: Logistic regression results on total sample, no interaction between birth weight and parity\textsuperscript{a} (n=3816)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.79</td>
<td>0.31</td>
<td>&lt;.0001</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Birth weight\textsuperscript{b}</td>
<td>-0.0013</td>
<td>0.0003</td>
<td>&lt;.0001</td>
<td>0.88</td>
<td>0.84, 0.93</td>
</tr>
<tr>
<td>African American</td>
<td>0.40</td>
<td>0.13</td>
<td>0.002</td>
<td>1.50</td>
<td>1.16, 1.94</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>-0.89</td>
<td>0.40</td>
<td>0.02</td>
<td>0.41</td>
<td>0.19, 0.89</td>
</tr>
<tr>
<td>PDA\textsuperscript{c} No treatment</td>
<td>0.70</td>
<td>0.21</td>
<td>0.0008</td>
<td>2.02</td>
<td>1.34, 3.04</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>0.10</td>
<td>0.17</td>
<td>0.56</td>
<td>1.10</td>
<td>0.80, 1.52</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.02</td>
<td>0.23</td>
<td>&lt;.0001</td>
<td>2.78</td>
<td>1.75, 4.38</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>0.28</td>
<td>0.14</td>
<td>0.04</td>
<td>1.33</td>
<td>1.01, 1.75</td>
</tr>
</tbody>
</table>

\textbf{AUC} = 0.685 (95% CI .654 - .717)

Goodness-of-fit statistics: Hosmer – Lemeshow p = 0.67; Deviance p=1.0 (Value/DF = 0.56);

Pearson p=0.50 (Value/DF = 0.99)

\textsuperscript{a}Data presented as multivariable odds ratio (95% confidence interval).

\textsuperscript{b}For every 100 gram increase in birth weight, the likelihood of NEC decreases by 0.88.

\textsuperscript{c}PDA = patent ductus arteriosus.
Figure 13: Predicted Probabilities from Logistic Regression Modeling for infants with and without NEC*

* Infants with NEC had significantly higher mean predicted probabilities compared to infants without NEC (9.6% vs. 6.7%, p<.0001).
Figure 14: Area under the ROC curve (95% confidence interval) for the final prediction model (n=3816)

AUC = area under the curve

Prediction Model N= 3816
AUC = 0.689 (.658 - .720)
Table 17: Comparison of Maternal/Pregnancy Characteristics by Birth Weight Category

<table>
<thead>
<tr>
<th></th>
<th>&lt; 1000 g</th>
<th>1000-1500g</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=1595</td>
<td>N=2221</td>
<td></td>
</tr>
<tr>
<td>Maternal age</td>
<td>25.9 (6.6)</td>
<td>26.3 (6.6)</td>
<td>0.07</td>
</tr>
<tr>
<td>Maternal age &lt; 21 y</td>
<td>852 (25)</td>
<td>92 (24)</td>
<td>0.78</td>
</tr>
<tr>
<td>Marital status – single (67 missing)</td>
<td>813 (52)</td>
<td>1026 (47)</td>
<td>0.004</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>1496 (94)</td>
<td>2096 (94)</td>
<td>0.43</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>982 (62)</td>
<td>1380 (62)</td>
<td>0.79</td>
</tr>
<tr>
<td>Diabetes/insulin dependentb</td>
<td>36 (4)</td>
<td>74 (5)</td>
<td>0.05</td>
</tr>
<tr>
<td>Hypertension</td>
<td>378 (24)</td>
<td>603 (27)</td>
<td>0.018</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>368 (23)</td>
<td>384 (17)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>258 (16)</td>
<td>353 (16)</td>
<td>0.81</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>900 (56)</td>
<td>1154 (52)</td>
<td>0.007</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>1134 (71)</td>
<td>1582 (71)</td>
<td>0.99</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>988 (62)</td>
<td>1386 (62)</td>
<td>0.77</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>960 (60)</td>
<td>1441 (65)</td>
<td>0.003</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>406 (25)</td>
<td>685 (31)</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

bCategorical variables reported as n (%) and continuous variables reported as mean (SD).

bMissing data on 1399 subjects (37% of data).
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>&lt; 1000 g</th>
<th>1000-1500g</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=1595</td>
<td>1110 (50)</td>
<td>662 (30)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Gender – Male</td>
<td>761 (48)</td>
<td>1110 (50)</td>
<td>0.17</td>
</tr>
<tr>
<td>Race – Black</td>
<td>593 (37)</td>
<td>662 (30)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>26.1 (2.05)</td>
<td>29.7 (2.04)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>SGA</td>
<td>316 (20)</td>
<td>545 (25)</td>
<td>0.001</td>
</tr>
<tr>
<td>Onset of NEC</td>
<td>154 (9.7)</td>
<td>109 (4.9)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Age at onset of NEC in days</td>
<td>21.8 (14.4)</td>
<td>14.3 (12.4)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score</td>
<td>6.2 (2.3)</td>
<td>7.7 (1.7)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>726 (46)</td>
<td>409 (18)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of life</td>
<td>884 (55)</td>
<td>450 (20)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Number of days on ventilator first week of lifeb</td>
<td>3.7 (3.5)</td>
<td>1.3 (2.3)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>911 (57)</td>
<td>683 (31)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>772 (48)</td>
<td>303 (14)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus (PDA)</td>
<td>876 (55)</td>
<td>446 (20)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Treatmentc</td>
<td>735 (46)</td>
<td>311 (14)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>126 (8)</td>
<td>38 (2)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>1185 (74)</td>
<td>821 (37)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>1019 (64)</td>
<td>436 (20)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Required O₂ at 6 hrs thru 24 hrs of ageg</td>
<td>1231 (83)</td>
<td>1042 (51)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of aged</td>
<td>1438 (97)</td>
<td>1699 (83)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrsd</td>
<td>1332 (90)</td>
<td>1085 (53)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrsd</td>
<td>1340 (90)</td>
<td>1246 (61)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>&lt; 1000 g</td>
<td>1000-1500g</td>
<td>P-value</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>----------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>N=1595</td>
<td>N=2221</td>
<td></td>
</tr>
<tr>
<td>High FIO2 (&gt;0.9)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>63 (6)</td>
<td>23 (2)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Low temperature (&lt;35°)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>165 (38)</td>
<td>94 (14)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Received 1&lt;sup&gt;st&lt;/sup&gt; enteral feed ≤ 3 days of age</td>
<td>645 (40)</td>
<td>1677 (76)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Achieved full feeds during NICU stay</td>
<td>3098 (90)</td>
<td>344 (90)</td>
<td>0.97</td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds&lt;sup&gt;f&lt;/sup&gt;</td>
<td>20.8 (13.3)</td>
<td>10.9 (8.0)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 5 days of age</td>
<td>18 (1)</td>
<td>349 (16)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 7 days of age</td>
<td>67 (4)</td>
<td>824 (37)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 14 days of age</td>
<td>496 (31)</td>
<td>1718 (77)</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<sup>a</sup>Categorical variables reported as n (%) and continuous variables reported as mean (SD).

<sup>b</sup>For years <1998, on conventional ventilator for first 10 days of life; for ≥1998, on conventional ventilator for first 7 days of life.

<sup>c</sup>PDA treatment with indomethacin and/or surgery.

<sup>d</sup>Missing data on 298 subjects (8% of data).

<sup>e</sup>Missing data FIO2 data on 1515 subjects (40% of data) and missing temperature data on 2713 subjects (71% of data).

<sup>f</sup>Among infants who achieved full enteral feeding during NICU stay.
<table>
<thead>
<tr>
<th></th>
<th>&lt;1000 n=1595</th>
<th></th>
<th></th>
<th>&lt;1000 n=1595</th>
<th></th>
<th></th>
<th>1000-1500 n=2221</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NEC</td>
<td>No NEC</td>
<td>OR or mean difference (SE)</td>
<td>95% CI</td>
<td>p</td>
<td>NEC</td>
<td>No NEC</td>
<td>OR or mean difference (SE)</td>
<td>95% CI</td>
</tr>
<tr>
<td>Maternal age</td>
<td>26.3 (6.6)</td>
<td>25.9 (6.6)</td>
<td>-.43 (.56)</td>
<td>-.15, .67</td>
<td>.444</td>
<td>25.1 (6.7)</td>
<td>26.3 (6.6)</td>
<td>1.3 (.64)</td>
<td>.01, 2.54</td>
</tr>
<tr>
<td>Maternal age &lt; 21 yrs</td>
<td>36 (23)</td>
<td>385 (27)</td>
<td>.84</td>
<td>.57, 1.24</td>
<td>.371</td>
<td>36 (33)</td>
<td>487 (23)</td>
<td>1.65</td>
<td>1.09, 2.48</td>
</tr>
<tr>
<td>Marital status – single</td>
<td>79 (52)</td>
<td>734 (52)</td>
<td>.99</td>
<td>.71, 1.39</td>
<td>.962</td>
<td>61 (57)</td>
<td>965 (47)</td>
<td>1.52</td>
<td>1.03, 2.25</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>147 (95)</td>
<td>1349 (94)</td>
<td>1.42</td>
<td>.64, 3.11</td>
<td>.384</td>
<td>103 (95)</td>
<td>1993 (94)</td>
<td>1.02</td>
<td>.44, 2.37</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>94 (61)</td>
<td>888 (62)</td>
<td>.99</td>
<td>.70, 1.39</td>
<td>.931</td>
<td>60 (55)</td>
<td>1320 (63)</td>
<td>.74</td>
<td>.50, 1.08</td>
</tr>
<tr>
<td>Diabetes/insulin dependentb</td>
<td>4 (4)</td>
<td>32 (4)</td>
<td>1.10</td>
<td>.38, 3.17</td>
<td>.862</td>
<td>6 (9)</td>
<td>68 (5)</td>
<td>1.76</td>
<td>.73, 4.20</td>
</tr>
<tr>
<td>Hypertension</td>
<td>39 (25)</td>
<td>339 (24)</td>
<td>1.09</td>
<td>.75, 1.61</td>
<td>.644</td>
<td>15 (14)</td>
<td>588 (28)</td>
<td>.41</td>
<td>.24, .72</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>33 (21)</td>
<td>335 (23)</td>
<td>.90</td>
<td>.60, 1.34</td>
<td>.598</td>
<td>21 (19)</td>
<td>363 (17)</td>
<td>1.15</td>
<td>.70, 1.87</td>
</tr>
<tr>
<td>PROMc</td>
<td>32 (21)</td>
<td>226 (16)</td>
<td>1.41</td>
<td>.93, 2.13</td>
<td>.103</td>
<td>23 (21)</td>
<td>330 (16)</td>
<td>1.44</td>
<td>.90, 2.32</td>
</tr>
<tr>
<td>Maternal antibiotic exposure</td>
<td>95 (62)</td>
<td>805 (56)</td>
<td>1.27</td>
<td>.90, 1.79</td>
<td>.169</td>
<td>63 (58)</td>
<td>1091 (52)</td>
<td>1.28</td>
<td>.87, 1.89</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>124 (81)</td>
<td>1104 (70)</td>
<td>1.75</td>
<td>1.15,2.65</td>
<td>.008</td>
<td>76 (70)</td>
<td>1506 (71)</td>
<td>.92</td>
<td>.61, 1.40</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>56 (36)</td>
<td>551 (38)</td>
<td>.93</td>
<td>.65, 1.30</td>
<td>.649</td>
<td>39 (36)</td>
<td>796 (38)</td>
<td>.92</td>
<td>.62, 1.38</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>106 (69)</td>
<td>854 (59)</td>
<td>1.52</td>
<td>1.06, 2.17</td>
<td>.021</td>
<td>73 (67)</td>
<td>1368 (65)</td>
<td>1.10</td>
<td>.73, 1.66</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>42 (27)</td>
<td>364 (25)</td>
<td>1.11</td>
<td>.76, 1.61</td>
<td>.586</td>
<td>33 (30)</td>
<td>652 (31)</td>
<td>.97</td>
<td>.64, 1.48</td>
</tr>
</tbody>
</table>

*Categorical variables reported as n (%) and continuous variables reported as mean (SD); p=p-value.
*bMissing values on 1399 subjects (n=590 infants < 1000g; n=809 infants > 1000g). cPROM = prolonged rupture of membranes.
<table>
<thead>
<tr>
<th>Predictor</th>
<th>&lt;1000g (n=1595)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NEC</td>
<td>No NEC</td>
<td>OR</td>
<td>95% CI</td>
<td>p</td>
<td>NEC</td>
<td>No NEC</td>
<td>OR</td>
<td>95% CI</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>Gender – Male</td>
<td>75 (49)</td>
<td>686 (48)</td>
<td>1.04</td>
<td>.75, 1.46</td>
<td>.802</td>
<td>64 (59)</td>
<td>1046 (50)</td>
<td>1.45</td>
<td>.98, 2.14</td>
<td>.064</td>
<td></td>
</tr>
<tr>
<td>Race – African American</td>
<td>65 (42)</td>
<td>528 (37)</td>
<td>1.27</td>
<td>.90, 1.77</td>
<td>.174</td>
<td>47 (43)</td>
<td>615 (29)</td>
<td>1.85</td>
<td>1.25, 2.73</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>25.8 (1.8)</td>
<td>26.1 (2.1)</td>
<td>.33</td>
<td>-.1, .67</td>
<td>.057</td>
<td>28.6 (1.6)</td>
<td>29.8 (2.0)</td>
<td>1.2 (.20)</td>
<td>.80, 1.58</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>746 (137)</td>
<td>778 (140)</td>
<td>32 (12)</td>
<td>9, 55</td>
<td>.007</td>
<td>1201 (149)</td>
<td>1264 (145)</td>
<td>63 (14)</td>
<td>35, 91</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>SGA</td>
<td>31 (20)</td>
<td>285 (20)</td>
<td>1.02</td>
<td>.68, 1.55</td>
<td>.917</td>
<td>12 (11)</td>
<td>533 (25)</td>
<td>.37</td>
<td>.20, .67</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>5-minute Apgar Score</td>
<td>6.2 (2.2)</td>
<td>6.2 (2.3)</td>
<td>.006 (.19)</td>
<td>-.38, .39</td>
<td>.98</td>
<td>7.6(1.7)</td>
<td>7.7 (1.7)</td>
<td>.10 (.17)</td>
<td>-.23, .42</td>
<td>.550</td>
<td></td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>67 (44)</td>
<td>659 (46)</td>
<td>.91</td>
<td>.65, 1.28</td>
<td>.598</td>
<td>23 (21)</td>
<td>386 (18)</td>
<td>1.20</td>
<td>.75, 1.92</td>
<td>.458</td>
<td></td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of life</td>
<td>95 (62)</td>
<td>789 (55)</td>
<td>1.33</td>
<td>.95, 1.87</td>
<td>.099</td>
<td>28 (26)</td>
<td>422 (20)</td>
<td>1.38</td>
<td>.89, 2.16</td>
<td>.148</td>
<td></td>
</tr>
<tr>
<td>Number of days on ventilator first week of life</td>
<td>3.9 (3.6)</td>
<td>3.7 (3.5)</td>
<td>-.19 (.30)</td>
<td>-.78, .40</td>
<td>.524</td>
<td>1.69 (2.1)</td>
<td>1.32 (2.3)</td>
<td>-.37 (.22)</td>
<td>-.81, .07</td>
<td>.097</td>
<td></td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>96 (62)</td>
<td>815 (57)</td>
<td>1.27</td>
<td>.90, 1.79</td>
<td>.171</td>
<td>45 (41)</td>
<td>638 (30)</td>
<td>1.62</td>
<td>1.10, 2.41</td>
<td>.015</td>
<td></td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>78 (51)</td>
<td>694 (48)</td>
<td>1.10</td>
<td>.79, 1.54</td>
<td>.557</td>
<td>18 (17)</td>
<td>285 (13)</td>
<td>1.27</td>
<td>.75, 2.13</td>
<td>.371</td>
<td></td>
</tr>
<tr>
<td>Predictor</td>
<td>&lt;1000 g (n=1595)</td>
<td>1000-1500 g (n=2221)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NEC</td>
<td>No NEC</td>
<td>OR</td>
<td>95% CI</td>
<td>p</td>
<td>NEC</td>
<td>No NEC</td>
<td>OR</td>
<td>95% CI</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>Patent Ductus Arteriosus</td>
<td>93 (60)</td>
<td>783 (54)</td>
<td>1.28</td>
<td>.91, 1.80</td>
<td>.152</td>
<td>40 (37)</td>
<td>406 (19)</td>
<td>2.44</td>
<td>1.63, 3.65</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>PDA Treatment</td>
<td>76 (49)</td>
<td>659 (46)</td>
<td>1.16</td>
<td>.83, 1.61</td>
<td>.392</td>
<td>23 (21)</td>
<td>288 (14)</td>
<td>1.69</td>
<td>1.05, 2.73</td>
<td>.029</td>
<td></td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>19 (12)</td>
<td>107 (7)</td>
<td>1.75</td>
<td>1.04, 2.95</td>
<td>.032</td>
<td>10 (9)</td>
<td>28 (1)</td>
<td>7.52</td>
<td>3.55, 15.91</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Surfactant use</td>
<td>108 (70)</td>
<td>1077 (75)</td>
<td>.79</td>
<td>.55, 1.14</td>
<td>.214</td>
<td>40 (37)</td>
<td>781 (37)</td>
<td>.99</td>
<td>.66, 1.47</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>103 (67)</td>
<td>916 (64)</td>
<td>1.15</td>
<td>.81, 1.63</td>
<td>.441</td>
<td>31 (29)</td>
<td>405 (19)</td>
<td>1.69</td>
<td>1.10, 2.60</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>Required O₂ at 6 hrs thru 24 hrs of age</td>
<td>122 (82)</td>
<td>1109 (83)</td>
<td>.91</td>
<td>.59, 1.42</td>
<td>.685</td>
<td>52 (53)</td>
<td>990 (51)</td>
<td>1.06</td>
<td>.71, 1.59</td>
<td>.784</td>
<td></td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of age</td>
<td>145 (97)</td>
<td>1293 (97)</td>
<td>1.12</td>
<td>.40, 3.18</td>
<td>.829</td>
<td>92 (93)</td>
<td>1607 (83)</td>
<td>2.70</td>
<td>1.24, 5.87</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>Need for resp. support at 24 hrs</td>
<td>134 (90)</td>
<td>1198 (90)</td>
<td>1.01</td>
<td>.57, 1.77</td>
<td>.982</td>
<td>53 (54)</td>
<td>1032 (53)</td>
<td>1.01</td>
<td>.67, 1.51</td>
<td>.960</td>
<td></td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrs</td>
<td>134 (90)</td>
<td>1206 (90)</td>
<td>.95</td>
<td>.54, 1.65</td>
<td>.832</td>
<td>67 (68)</td>
<td>1179 (61)</td>
<td>1.35</td>
<td>.87, 2.07</td>
<td>.175</td>
<td></td>
</tr>
<tr>
<td>High FIO₂ (&gt;0.9)</td>
<td>11 (11)</td>
<td>523 (6)</td>
<td>2.02</td>
<td>1.02, 4.01</td>
<td>.041</td>
<td>1 (2)</td>
<td>22 (2)</td>
<td>.92</td>
<td>.12, 6.95</td>
<td>.937</td>
<td></td>
</tr>
<tr>
<td>Low temperature at admission (&lt;35°)</td>
<td>22 (49)</td>
<td>143 (36)</td>
<td>1.69</td>
<td>.91, 3.13</td>
<td>.096</td>
<td>2 (8)</td>
<td>92 (14)</td>
<td>.52</td>
<td>.12, 2.23</td>
<td>.367</td>
<td></td>
</tr>
<tr>
<td>Predictor</td>
<td>&lt;1000 g (n=1595)</td>
<td>1000-1500 g (n=2221)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NEC</td>
<td>No NEC</td>
<td>OR</td>
<td>95% CI</td>
<td>p</td>
<td>NEC</td>
<td>No NEC</td>
<td>OR</td>
<td>95% CI</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>Received 1st enteral feed ≤ 3 days of age</td>
<td>67 (44)</td>
<td>578 (40)</td>
<td>1.15</td>
<td>.82, 1.61</td>
<td>.415</td>
<td>84 (77)</td>
<td>1593 (75)</td>
<td>1.09</td>
<td>.69, 1.73</td>
<td>.698</td>
<td></td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds (among those achieving)</td>
<td>25.4 (17)</td>
<td>20.3 (13)</td>
<td>-5.0 (1.3)</td>
<td>-7.5, -2.6</td>
<td>.003</td>
<td>18.7 (15.3)</td>
<td>10.5 (7.3)</td>
<td>-8.2 (.82)</td>
<td>-.9.8, -6.6</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Full enteral feeds by 5 days of age</td>
<td>2 (1)</td>
<td>16 (1)</td>
<td>1.2</td>
<td>.27, 5.1</td>
<td>.833</td>
<td>5 (5)</td>
<td>344 (16)</td>
<td>.25</td>
<td>.10, .61</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Full enteral feeds by 7 days of age</td>
<td>4 (3)</td>
<td>63 (4)</td>
<td>.58</td>
<td>.21, 1.63</td>
<td>.297</td>
<td>25 (23)</td>
<td>799 (38)</td>
<td>.49</td>
<td>.31, .77</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Full enteral feeds by 14 days of age</td>
<td>40 (26)</td>
<td>456 (32)</td>
<td>.76</td>
<td>.52, 1.1</td>
<td>.149</td>
<td>54 (50)</td>
<td>1664 (79)</td>
<td>.26</td>
<td>.18, .39</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>

*aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

*bFor years <1998, on conventional ventilator for first 10 days of life; for ≥1998, on conventional ventilator for first 7 days of life.

*cPDA (patent ductus arteriosus) treatment with indomethacin and/or surgery.

*d298 missing values (n=113 infants < 1000g; n=185 infants ≥ 1000g).

*e1515 missing FIO2 values (n=615 infants < 1000g; n=900 infants ≥ 1000g); 2713 missing temperature values (n=1155 infants <1000g; n=1558 infants ≥ 1000g).
Table 21: Stratified Multivariate Logistic Regression Results: Infants 401-999 g birth weight

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.757</td>
<td>0.5033</td>
<td>0.0005</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Birth weight&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.002</td>
<td>0.0006</td>
<td>0.006</td>
<td>0.92</td>
<td>0.87, 0.98</td>
</tr>
<tr>
<td>Surgical PDA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.613</td>
<td>0.268</td>
<td>0.022</td>
<td>1.85</td>
<td>1.09, 3.12</td>
</tr>
<tr>
<td>Parity of child &gt; 1</td>
<td>0.459</td>
<td>0.184</td>
<td>0.012</td>
<td>1.58</td>
<td>1.11, 2.27</td>
</tr>
<tr>
<td>Antenatal Steroids</td>
<td>0.589</td>
<td>0.213</td>
<td>0.006</td>
<td>1.80</td>
<td>1.19, 2.74</td>
</tr>
</tbody>
</table>

\[ \text{AUC} = 0.61 \ (0.56, \ 0.66) \]

Goodness-of-fit statistics: Hosmer-Lemeshow p=.12; Deviance p=1.0 ; Pearson p=0.45

<sup>a</sup>For every 50 gram increase in birth weight.

<sup>b</sup>Patent Ductus Arteriosus.
### Table 22: Stratified Multivariate Logistic Regression Results: Infants 1000-1500g birth weight

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\beta$</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>-0.59</td>
<td>0.91</td>
<td>0.52</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Birth weight$^a$</td>
<td>-0.002</td>
<td>0.0007</td>
<td>0.003</td>
<td>0.90</td>
<td>0.84, 0.96</td>
</tr>
<tr>
<td>African American</td>
<td>0.67</td>
<td>0.21</td>
<td>0.002</td>
<td>1.95</td>
<td>1.29, 2.94</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeds in first 5 days of life</td>
<td>-1.06</td>
<td>0.47</td>
<td>0.025</td>
<td>0.35</td>
<td>0.14, 0.89</td>
</tr>
<tr>
<td>PDA$^b$ No treatment</td>
<td>1.09</td>
<td>0.30</td>
<td>0.0002</td>
<td>2.98</td>
<td>1.67, 5.31</td>
</tr>
<tr>
<td>PDA$^b$ Indomethacin only</td>
<td>-0.05</td>
<td>0.32</td>
<td>0.88</td>
<td>0.95</td>
<td>0.51, 1.79</td>
</tr>
<tr>
<td>PDA$^b$ Surgery</td>
<td>2.01</td>
<td>0.41</td>
<td>&lt;.0001</td>
<td>7.49</td>
<td>3.36, 16.67</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Maternal age &lt; 21y</td>
<td>0.40</td>
<td>0.22</td>
<td>0.066</td>
<td>1.51</td>
<td>0.98, 2.95</td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>-0.72</td>
<td>0.29</td>
<td>0.013</td>
<td>0.49</td>
<td>0.28, 0.86</td>
</tr>
</tbody>
</table>

$AUC = 0.730$ (95% CI 0.682, 0.780)

Goodness-of-fit statistics: Hosmer-Lemeshow $p=0.42$; Deviance $p=1.0$ (Value/DF 0.41);
Pearson $p=0.99$ (Value/DF 0.91)

$^a$Odds Ratio for every 50 gram increase in birth weight.

$^b$PDA=patent ductus arteriosus.
Figure 15: Interaction between Birth Weight and Maternal Hypertension
Figure 16: Area Under the Receiver Operator Curves (AUC) for 2 Multivariable Logistic Regression Models. Modeling the likelihood of NEC, stratified by birth weight.

- **Prediction Model for infants 1000-1500 g:**
  - AUC = 0.73 (95% CI 0.69, 0.78)

- **Prediction Model for infants 401-999 g:**
  - AUC = 0.61 (95% CI 0.56, 0.66)
*For both groups of infants (< 1000 g and 1000-1500 g), the mean predicted probabilities between infants with and without NEC was significantly different (p<.0001).
### Table 23: Logistic regression results on total sample with hypertension included (n=3816)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predictor</strong></td>
<td>β</td>
<td>SE</td>
<td>p-value</td>
<td>Odds Ratio</td>
<td>95% CI</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.50</td>
<td>0.46</td>
<td>&lt;.0001</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.0005</td>
<td>0.0004</td>
<td>0.29</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0.42</td>
<td>0.13</td>
<td>0.002</td>
<td>1.52</td>
<td>1.17, 1.97</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>-0.90</td>
<td>0.40</td>
<td>0.02</td>
<td>0.41</td>
<td>0.19, 0.89</td>
</tr>
<tr>
<td>PDA(^b) No treatment</td>
<td>0.67</td>
<td>0.21</td>
<td>0.001</td>
<td>1.97</td>
<td>1.30, 2.97</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>0.09</td>
<td>0.17</td>
<td>0.61</td>
<td>1.09</td>
<td>0.79, 1.51</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.01</td>
<td>0.24</td>
<td>&lt;.0001</td>
<td>2.76</td>
<td>1.73, 4.39</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>1.42</td>
<td>0.50</td>
<td>0.005</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>Parity x birth weight(^a)</td>
<td>-0.0012</td>
<td>0.0005</td>
<td>0.016</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing weight</td>
<td>---</td>
<td></td>
<td>&lt;.0001</td>
<td>1.18</td>
<td>1.11, 1.26</td>
</tr>
<tr>
<td>Parity = 1, decreasing weight</td>
<td>---</td>
<td></td>
<td>0.29</td>
<td>1.05</td>
<td>0.96, 1.14</td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>-0.17</td>
<td>0.16</td>
<td>0.29</td>
<td>0.84</td>
<td>0.61, 1.16</td>
</tr>
</tbody>
</table>

**AUC = 0.693 (95% CI 0.662 – 0.724)**

Goodness-of-fit statistics: Hosmer – Lemeshow p = 0.13
Deviance p=1.0 (Value/DF = 0.53), Pearson p=0.98 (Value/DF = 0.95)

\(^a\)Among infants who are not the first born (parity > 1), for every 100 gram decrease in birth weight, the likelihood of NEC increases by 1.18.

\(^b\)PDA = patent ductus arteriosus.
**Table 24: Logistic regression results on total sample after adding hypertension and birth weight interaction term (n=3816)**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>SE</th>
<th>p-value</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.89</td>
<td>0.49</td>
<td>&lt;.0001</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.00004</td>
<td>0.0005</td>
<td>0.94</td>
<td>------------</td>
<td>1.0</td>
</tr>
<tr>
<td>African American</td>
<td>0.41</td>
<td>0.13</td>
<td>0.002</td>
<td>1.51</td>
<td>1.17, 1.97</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>-0.89</td>
<td>0.40</td>
<td>0.026</td>
<td>0.41</td>
<td>0.19, 0.90</td>
</tr>
<tr>
<td>PDA No treatment</td>
<td>0.68</td>
<td>0.21</td>
<td>0.001</td>
<td>1.97</td>
<td>1.30, 2.98</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>0.09</td>
<td>0.17</td>
<td>0.60</td>
<td>1.09</td>
<td>0.79, 1.51</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.04</td>
<td>0.24</td>
<td>&lt;.0001</td>
<td>2.82</td>
<td>1.78, 4.48</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>1.55</td>
<td>0.50</td>
<td>0.002</td>
<td>------------</td>
<td>1.0</td>
</tr>
<tr>
<td>Parity x birth weight</td>
<td>-0.0014</td>
<td>0.0005</td>
<td>0.007</td>
<td>------------</td>
<td>1.15</td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing weight</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
<td>1.15</td>
<td>1.08, 1.23</td>
</tr>
<tr>
<td>Parity = 1, decreasing weight</td>
<td></td>
<td></td>
<td>0.94</td>
<td>1.00</td>
<td>0.92, 1.10</td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>1.09</td>
<td>0.55</td>
<td>0.049</td>
<td>------------</td>
<td>1.15</td>
</tr>
<tr>
<td>Hypertension x birth weight</td>
<td>-0.0014</td>
<td>0.0006</td>
<td>0.02</td>
<td>------------</td>
<td>1.15</td>
</tr>
<tr>
<td>Hypertension = yes, decreasing weight</td>
<td></td>
<td></td>
<td>0.019</td>
<td>1.15</td>
<td>1.02, 1.29</td>
</tr>
<tr>
<td>Hypertension = no, decreasing weight</td>
<td></td>
<td></td>
<td>0.94</td>
<td>1.00</td>
<td>0.92, 1.10</td>
</tr>
</tbody>
</table>

*AUC = 0.694 (95% CI 0.664 – 0.726)*

Goodness-of-fit statistics: Hosmer – Lemeshow p = 0.31
Deviance p=1.0 (Value/DF 0.53), Pearson p=0.98 (Value/DF 0.95)

*a*PDA = patent ductus arteriosus.

*b*Among infants who are not the first born (parity > 1), for every 100 gram decrease in birth weight, the likelihood of NEC increases by 1.15.

*c*Among infants whose mother had hypertension during pregnancy, for every 100 gram decrease in birth weight, the likelihood of NEC increases by 1.15.
Table 25: Comparisons of the area under the ROC curve and AIC candidate multivariable logistic models

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables</th>
<th>N</th>
<th>AUC</th>
<th>AIC</th>
<th>-2 log likelihood</th>
<th>p-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BTWT, AA, Feeds, PDA, Parity</td>
<td>3816</td>
<td>0.685</td>
<td>1831.756</td>
<td>1815.756</td>
<td>0.017&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.654 – .717)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>BTWT, AA, Feeds, PDA, Parity, BTWT x Parity</td>
<td>3816</td>
<td>0.689</td>
<td>1828.069</td>
<td>1810.069</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.658 – .720)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>BTWT, AA, Feeds, PDA, Parity, BTWT x Parity,</td>
<td>3801</td>
<td>0.693</td>
<td>1826.319</td>
<td>1806.319</td>
<td>0.019&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Hypertension</td>
<td></td>
<td>(0.662 - .724)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>BTWT, AA, Feeds, PDA, Parity, BTWT x Parity,</td>
<td>3801</td>
<td>0.694</td>
<td>1822.807</td>
<td>1800.807</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BTWT x Hypertension</td>
<td></td>
<td>(0.664 – .726)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Based on the differences of the -2 log likelihoods of 2 models. DF = degrees of freedom; AIC = Akaike’s Information Criterion; ROC = receiver operator characteristic; CI=confidence interval; BTWT = birth weight.

<sup>b</sup>The addition of interaction term (BTWT x Parity) in Model 2 considered statistically significant, ($\chi^2$ of 5.69 with 1 DF).

<sup>c</sup>The addition of interaction term (BTWT x Hypertension) in Model 4 considered statistically significant, ($\chi^2$ of 5.51 with 1 DF). However, the addition of the variable for Hypertension to the Model 2 is only marginally significant (Comparison of Model 2 and Model 3: $\chi^2$ of 3.75, p=0.053).
Table 26: Comparisons of the area under the ROC curve (AUC) for all univariable models compared to the final multivariable prediction model\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariable model(^b)</td>
<td>.695</td>
<td>.664, .726</td>
</tr>
<tr>
<td>Birth weight</td>
<td>.636</td>
<td>.603, .669</td>
</tr>
<tr>
<td>PDA categories</td>
<td>.600</td>
<td>.579, .621</td>
</tr>
<tr>
<td>African American</td>
<td>.552</td>
<td>.536, .568</td>
</tr>
<tr>
<td>Reach full enteral feeds by 5 days</td>
<td>.537</td>
<td>.478, .596</td>
</tr>
<tr>
<td>Parity</td>
<td>.528</td>
<td>.487, .569</td>
</tr>
<tr>
<td>Hypertension</td>
<td>.528</td>
<td>.503, .554</td>
</tr>
</tbody>
</table>

\(^a\)All comparisons with multivariable model statistically significant, p<0.001; ROC = receiver operator characteristic; CI=confidence interval.

\(^b\)Multivariable model contains birth weight, race, feedings, PDA, parity, parity x birth weight interaction, hypertension, hypertension x birth weight interaction.
Table 27: Maternal Characteristics according to Parity of Child\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>Parity &gt; 1</th>
<th>Parity = 1</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=2401</td>
<td>N=1415</td>
<td></td>
</tr>
<tr>
<td>Maternal age</td>
<td>27.7 (6.2)</td>
<td>23.4 (6.2)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Maternal age &lt; 21 yrs</td>
<td>337 (14)</td>
<td>607 (43)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Marital status – single</td>
<td>1024 (43)</td>
<td>815 (59)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>2229 (93)</td>
<td>1363 (96)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>1558 (65)</td>
<td>804 (57)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diabetes/insulin dependent(^b)</td>
<td>74 (5)</td>
<td>36 (4)</td>
<td>0.36</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>500 (21)</td>
<td>252 (18)</td>
<td>0.025</td>
</tr>
<tr>
<td>Hypertension</td>
<td>485 (20)</td>
<td>496 (35)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>408 (17)</td>
<td>203 (14)</td>
<td>0.03</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>1300 (54)</td>
<td>754 (53)</td>
<td>0.63</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>1721 (72)</td>
<td>995 (70)</td>
<td>0.36</td>
</tr>
</tbody>
</table>

\(^a\)Categorical variables reported as n (%) and continuous variables reported as mean (SD).

\(^b\)Missing values on 1399 subjects (37% of the data).
Table 28: Clinical characteristics of VLBW infants according to Parity of Childa

<table>
<thead>
<tr>
<th></th>
<th>Parity&gt;1 N=2400</th>
<th>Parity=1 N=1415</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender – Male</td>
<td>1167 (49)</td>
<td>704 (50)</td>
<td>0.50</td>
</tr>
<tr>
<td>Race – African American</td>
<td>807 (34)</td>
<td>448 (32)</td>
<td>0.22</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>28.2 (2.7)</td>
<td>28.2 (2.8)</td>
<td>0.73</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>1068 (276)</td>
<td>1040 (284)</td>
<td>0.002</td>
</tr>
<tr>
<td>SGA</td>
<td>489 (20)</td>
<td>372 (26)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score – Mean (SD)</td>
<td>7.1 (2.1)</td>
<td>7.0 (2.1)</td>
<td>0.25</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>698 (29)</td>
<td>437 (31)</td>
<td>0.24</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of lifeb</td>
<td>832 (35)</td>
<td>502 (35)</td>
<td>0.61</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>1011 (42)</td>
<td>583 (41)</td>
<td>0.59</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>648 (27)</td>
<td>427 (30)</td>
<td>0.03</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus</td>
<td>823 (34)</td>
<td>499 (35)</td>
<td>0.54</td>
</tr>
<tr>
<td>PDA Treatmentc</td>
<td>635 (26)</td>
<td>411 (29)</td>
<td>0.08</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>95 (4)</td>
<td>69 (5)</td>
<td>0.18</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>1273 (53)</td>
<td>733 (52)</td>
<td>0.48</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>887 (37)</td>
<td>568 (40)</td>
<td>0.05</td>
</tr>
<tr>
<td>Required O₂ at 6 hrs thru 24 hrs of aged</td>
<td>1409 (64)</td>
<td>864 (66)</td>
<td>0.13</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of aged</td>
<td>1963 (89)</td>
<td>1174 (90)</td>
<td>0.25</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrsd</td>
<td>1517 (69)</td>
<td>900 (69)</td>
<td>0.80</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrsd</td>
<td>1614 (73)</td>
<td>972 (74)</td>
<td>0.31</td>
</tr>
<tr>
<td>Received 1st enteral feed ≤ 3 days of age</td>
<td>1489 (62)</td>
<td>833 (59)</td>
<td>0.05</td>
</tr>
<tr>
<td>Full enteral feeds by 5 days of age</td>
<td>226 (9)</td>
<td>141 (10)</td>
<td>0.58</td>
</tr>
</tbody>
</table>

aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

bFor years <1998, on conventional ventilator for first 10 days of life; for ≥1998, on conventional ventilator for first 7 days of life.

cPDA (patent ductus arteriosus) treatment with indomethacin and/or surgery.

dMissing data on 298 subjects (8% of data).
Figure 18: Interaction between Birth Weight and Parity of infant

![Graph showing the interaction between birth weight and parity. The x-axis represents birth weight, and the y-axis represents predicted logits. Two lines are depicted: one for Parity = 1: First born child and another for Parity > 1. The graph illustrates that as birth weight increases, the predicted logits decrease for both parity levels.]
Table 29: Maternal characteristics according to pregnancy induced hypertension status

<table>
<thead>
<tr>
<th></th>
<th>Hypertension N=981</th>
<th>No Hypertension N=2820</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>27.0 (6.7)</td>
<td>25.8 (6.5)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Maternal age &lt; 21 yrs</td>
<td>212 (22)</td>
<td>727 (26)</td>
<td>0.009</td>
</tr>
<tr>
<td>Marital status – single</td>
<td>417 (43)</td>
<td>1413 (51)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>963 (98)</td>
<td>2626 (93)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>836 (85)</td>
<td>1517 (54)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diabetes/insulin dependentb</td>
<td>36 (5)</td>
<td>74 (4)</td>
<td>0.19</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>75 (8)</td>
<td>674 (24)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>30 (3)</td>
<td>577 (20)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>300 (31)</td>
<td>1749 (62)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>782 (80)</td>
<td>1932 (69)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>95 (36)</td>
<td>1347 (38)</td>
<td>0.56</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>485 (49)</td>
<td>1907 (68)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>178 (18)</td>
<td>913 (32)</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

bMissing values on 1399 subjects (37% of the data).
Table 30: Clinical characteristics of VLBW infants by pregnancy induced hypertension status

<table>
<thead>
<tr>
<th></th>
<th>Hypertension</th>
<th>Hypertension</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=981</td>
<td>N=2820</td>
<td></td>
</tr>
<tr>
<td>Gender – Male</td>
<td>436 (22)</td>
<td>1427 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Race – African American</td>
<td>312 (32)</td>
<td>933 (33)</td>
<td>0.46</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>29.6 (2.4)</td>
<td>27.7 (2.7)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>1077 (282)</td>
<td>1051 (278)</td>
<td>0.013</td>
</tr>
<tr>
<td>SGA</td>
<td>470 (48)</td>
<td>390 (14)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score</td>
<td>7.4 (1.8)</td>
<td>6.9 (2.1)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>219 (22)</td>
<td>905 (32)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk of life(^b)</td>
<td>295 (30)</td>
<td>1039 (37)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>239 (24)</td>
<td>1345 (48)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>220 (22)</td>
<td>852 (30)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus</td>
<td>268 (27)</td>
<td>1049 (37)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Treatment(^c)</td>
<td>210 (21)</td>
<td>832 (29.5)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>16 (1.6)</td>
<td>147 (5)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>443 (45)</td>
<td>1553 (55)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>276 (28)</td>
<td>1170 (42)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Required O(_2) at 6 hrs thru 24 hrs of age(^d)</td>
<td>529 (58)</td>
<td>1742 (67)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of age(^e)</td>
<td>780 (85)</td>
<td>2355 (91)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrs(^d)</td>
<td>555 (60)</td>
<td>1860 (72)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrs(^d)</td>
<td>608 (66)</td>
<td>1976 (76)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Received 1st enteral feed &lt; 3 days of age</td>
<td>646 (66)</td>
<td>1670 (59)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Full enteral feeds by 5 days of age</td>
<td>123 (13)</td>
<td>243 (9)</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

\(^a\)Categorical variables reported as n (%) and continuous variables reported as mean (SD).

\(^b\)For years <1998, on conventional ventilator for first 10 days of life; for \(\geq\)1998, on conventional ventilator for first 7 days of life.

\(^c\)PDA (patent ductus arteriosus) treatment with indomethacin and/or surgery.

\(^d\)Missing data on 298 subjects (8% of data).
Table 31. Unadjusted associations between Maternal/Pregnancy Characteristics and NEC (N=39940)\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>NEC N=2820</th>
<th>No NEC N=37120</th>
<th>OR or mean difference (SE)</th>
<th>95% CI</th>
<th>(p^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>25.9 (6.7)</td>
<td>26.7 (6.8)</td>
<td>0.72 (0.13)</td>
<td>0.46-0.98</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Maternal age &lt; 21 yrs</td>
<td>713 (25)</td>
<td>8386 (23)</td>
<td>1.2</td>
<td>1.1-1.3</td>
<td>.001</td>
</tr>
<tr>
<td>Marital status – single</td>
<td>1590 (57)</td>
<td>19,247 (52)</td>
<td>1.2</td>
<td>1.1-1.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>At least one prenatal care visit</td>
<td>2567 (91)</td>
<td>33,975 (92)</td>
<td>0.94</td>
<td>0.82-1.1</td>
<td>0.34</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>1449 (52)</td>
<td>21,438 (58)</td>
<td>0.77</td>
<td>0.72-0.83</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diabetes/insulin dependent</td>
<td>86 (3.1)</td>
<td>1127 (3.1)</td>
<td>1.0</td>
<td>0.8-1.3</td>
<td>.98</td>
</tr>
<tr>
<td>Hypertension</td>
<td>510 (18)</td>
<td>9746 (27)</td>
<td>0.62</td>
<td>0.56-0.68</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>493 (18)</td>
<td>5727 (16)</td>
<td>1.2</td>
<td>1.1-1.3</td>
<td>.0036</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>708 (26)</td>
<td>7767 (21)</td>
<td>1.3</td>
<td>1.2-1.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>1752 (63)</td>
<td>21,057 (57)</td>
<td>1.3</td>
<td>1.2-1.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>1890 (68)</td>
<td>23,168 (64)</td>
<td>1.2</td>
<td>1.1-1.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>850 (30)</td>
<td>11860 (32)</td>
<td>0.92</td>
<td>0.85-0.999</td>
<td>0.047</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>1786 (63)</td>
<td>23,365 (63)</td>
<td>1.02</td>
<td>0.94-1.10</td>
<td>0.64</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>610 (22)</td>
<td>8963 (24)</td>
<td>0.87</td>
<td>0.79-0.95</td>
<td>.0026</td>
</tr>
</tbody>
</table>

\(^a\)Categorical variables reported as n (%) and continuous variables reported as mean (SD).

\(^b\)P-value from chi-square test (categorical variables) or t-test (continuous variables).

Missing information for maternal age: NEC n=2, no NEC n= 29; marital status 4, 47; prenatal care: 10, 136; c-section: 4, 77; diabetes: 17, 265; hypertension: 20, 287; hemorrhage: 14, 162; gravida: 2, 21; parity: 4, 23; multiple birth: 0, 1.
Table 32: Unadjusted associations between clinical characteristics of VLBW infants and NEC

<table>
<thead>
<tr>
<th></th>
<th>NEC N=2820</th>
<th>No NEC N=37120</th>
<th>OR or mean difference (SE)</th>
<th>95% CI</th>
<th>p^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender – Male</td>
<td>1551 (55)</td>
<td>18,546 (50)</td>
<td>1.2</td>
<td>1.1-1.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Race – African American</td>
<td>1431 (51)</td>
<td>15,674 (42)</td>
<td>1.4</td>
<td>1.3-1.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1431 (50.8)</td>
<td>15,674 (42)</td>
<td>1.4 REF</td>
<td>1.3-1.6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>White</td>
<td>941 (33.4)</td>
<td>14,717 (39.7)</td>
<td>1.03</td>
<td>0.9-1.2</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>353 (12.5)</td>
<td>5371 (14.5)</td>
<td>1.1</td>
<td>0.9-1.4</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>93 (3.3)</td>
<td>1275 (3.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>26.8 (2.5)</td>
<td>28.2 (2.8)</td>
<td>1.4 (.06)</td>
<td>1.3-1.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>917 (266)</td>
<td>1058 (283)</td>
<td>142 (5.5)</td>
<td>131-152</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>SGA</td>
<td>402 (14)</td>
<td>8167 (22)</td>
<td>0.6</td>
<td>0.5-0.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score</td>
<td>6.8 (1.9)</td>
<td>7.1 (1.8)</td>
<td>0.31 (0.04)</td>
<td>0.24-0.38</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>967 (35)</td>
<td>10,426 (28)</td>
<td>1.3</td>
<td>1.2-1.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mechanical ventilation &gt; 2 days in 1st wk of life</td>
<td>1641 (65)</td>
<td>15,082 (48)</td>
<td>2.0</td>
<td>1.9-2.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Number of days on ventilator first week of life</td>
<td>4.8 (3.4)</td>
<td>3.5 (3.4)</td>
<td>-1.246 (.07)</td>
<td>-1.4, -1.1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>1584 (58)</td>
<td>16,091 (44)</td>
<td>1.7</td>
<td>1.6-1.9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Indomethacin for any indication</td>
<td>779 (28)</td>
<td>6489 (18)</td>
<td>1.8</td>
<td>1.7-2.0</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Patent Ductus Arteriosus</td>
<td>1198 (42.5)</td>
<td>11,221 (30)</td>
<td>1.7</td>
<td>1.6-1.8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Treatment</td>
<td>972 (35)</td>
<td>9273 (25)</td>
<td>1.6</td>
<td>1.5-1.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>328 (12)</td>
<td>2203 (6)</td>
<td>2.1</td>
<td>1.8-2.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>1896 (67)</td>
<td>21,611 (58)</td>
<td>1.5</td>
<td>1.4-1.6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>2007 (71)</td>
<td>21,181 (57)</td>
<td>1.9</td>
<td>1.7-2.0</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Required O2 at 6 hrs thru 24 hrs of age</td>
<td>1544 (61)</td>
<td>17,525 (54)</td>
<td>1.4</td>
<td>1.3-1.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of age</td>
<td>2310 (91)</td>
<td>27,516 (84)</td>
<td>2.0</td>
<td>1.7-2.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrs</td>
<td>2093 (83)</td>
<td>23,242 (71)</td>
<td>2.0</td>
<td>1.8-2.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>NEC N=2820</td>
<td>No NEC N=37120</td>
<td>OR or mean difference (SE)</td>
<td>95% CI</td>
<td>p</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24 hrs</td>
<td>2088 (82)</td>
<td>24,010 (74)</td>
<td>1.7</td>
<td>1.6-1.9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>High FIO2 (&gt;0.9)</td>
<td>0.49 (0.35)</td>
<td>0.30 (0.24)</td>
<td>-0.19 (.01)</td>
<td>-0.2, -1.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Low temperature at admission (&lt;35°)</td>
<td>131 (18)</td>
<td>1278 (14)</td>
<td>1.4</td>
<td>1.1-1.7</td>
<td>.0025</td>
</tr>
<tr>
<td>Mean days to reach first enteral feeds (among those achieving)</td>
<td>7.9 (9.3)</td>
<td>5.6 (5.8)</td>
<td>-2.3 (0.1)</td>
<td>-2.6, -2.1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Received 1st enteral feed ≤ 3 d of age</td>
<td>838 (30)</td>
<td>14,539 (40)</td>
<td>0.65</td>
<td>0.59-0.70</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds (among those achieving)</td>
<td>31.5 (23.5)</td>
<td>18.6 (13.7)</td>
<td>-12.9 (0.3)</td>
<td>-13.5, -12.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Full enteral feeds by 5 days of age</td>
<td>38 (1.4)</td>
<td>2057 (5.7)</td>
<td>0.24</td>
<td>0.17-0.33</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

*aCategorical variables reported as n (%) and continuous variables reported as mean (SD).

bP-value from chi-square test (categorical variables) or t-test (continuous variables).

Missing information for gender: NEC group 0, no NEC group 1; race: 2, 83; gestational age: 3, 19; SGA: 3, 22; 5 minute apgar score: 30, 463; ventilation: 314, 5922; PROM: 97, 930; maternal antibiotics: 28, 179; infant antibiotic: 67, 528; indomethacin: 29, 178; PDA: 2, 11; PDA treatment: 3, 20; PDA surgery: 2, 15; antenatal steroids: 26, 162; surfactant use: 3, 39; intubation: 6, 59; oxygen 6-24 hours: 307, 4501; resp distress at 24 hours: 292, 4446; resp support at 24 hours: 293, 4458; abnormal chest x-ray at 24 hours: 299, 4508.

Missing information for enteral feeds: NEC group 23, no NEC group 586; full enteral feeds: 113, 885.
Figure 19: Calibration Plot for Multivariable Logistic Prediction Models: Internal Sample (N=3,816) and External Sample (n=32,770)*

*Predicted probabilities by observed probabilities for NEC. Hosmer-Lemeshow statistic (H-L).
Table 33: Logistic regression results—Internal Sample and External Validation Sample\(^a\)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model N=3816</th>
<th></th>
<th></th>
<th>P</th>
<th>External Model N=32,770</th>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\beta)</td>
<td>OR</td>
<td>95% CI</td>
<td>(P)</td>
<td>(\beta)</td>
<td>OR</td>
<td>95% CI</td>
<td>(P)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.44</td>
<td></td>
<td>&lt;.0001</td>
<td></td>
<td>-1.16</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight(^b)</td>
<td>-.00004</td>
<td></td>
<td>0.94</td>
<td></td>
<td>-0.0132</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0.42</td>
<td>1.51</td>
<td>1.17, 1.97</td>
<td>0.002</td>
<td>0.23</td>
<td>1.27</td>
<td>1.16, 1.37</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>-0.88</td>
<td>0.41</td>
<td>.19, .90</td>
<td>0.026</td>
<td>-0.80</td>
<td>0.45</td>
<td>0.32, 0.63</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA(^c) No treatment</td>
<td>0.68</td>
<td>1.97</td>
<td>1.3, 2.98</td>
<td>0.001</td>
<td>0.36</td>
<td>1.44</td>
<td>1.23, 1.68</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Indomethacin</td>
<td>0.09</td>
<td>1.09</td>
<td>0.79, 1.51</td>
<td>0.60</td>
<td>0.04</td>
<td>1.05</td>
<td>0.94, 1.16</td>
<td>0.19</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>1.04</td>
<td>2.82</td>
<td>1.78, 4.48</td>
<td>&lt;.0001</td>
<td>0.37</td>
<td>1.45</td>
<td>1.27, 1.66</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1</td>
<td>1.55</td>
<td></td>
<td></td>
<td>0.005</td>
<td>0.11</td>
<td></td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Parity x birth wt(^b)</td>
<td>-.0014</td>
<td></td>
<td></td>
<td>0.016</td>
<td>-0.0001</td>
<td></td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Parity &gt; 1, decreasing wt(^b)</td>
<td>--------</td>
<td>1.15</td>
<td>1.08, 1.23</td>
<td>&lt;.0001</td>
<td>--------</td>
<td>1.15</td>
<td>1.13, 1.17</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Parity = 1, decreasing weight(^b)</td>
<td>--------</td>
<td>1.00</td>
<td>0.92, 1.10</td>
<td>0.28</td>
<td>--------</td>
<td>1.14</td>
<td>1.11, 1.18</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hypertension(^d)</td>
<td>1.09</td>
<td></td>
<td></td>
<td>0.049</td>
<td>0.45</td>
<td></td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>Hypertension x btwt(^b)</td>
<td>-.002</td>
<td></td>
<td></td>
<td>0.02</td>
<td>-.001</td>
<td></td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Hypertension = yes, decreasing weight</td>
<td>--------</td>
<td>1.15</td>
<td>1.02, 1.29</td>
<td>0.019</td>
<td>--------</td>
<td>1.25</td>
<td>1.21, 1.29</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hypertension = no, decreasing weight</td>
<td>--------</td>
<td>1.00</td>
<td>0.92, 1.10</td>
<td>0.094</td>
<td>--------</td>
<td>1.14</td>
<td>1.11, 1.17</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td><strong>AUC</strong></td>
<td>0.695</td>
<td>(95% CI 0.664 – 0.726)</td>
<td></td>
<td>0.664 (95% CI 0.65 – 0.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)\(\beta\) = beta coefficient from logistic regression, OR = odds ratio, 95% CI = 95% confidence interval, \(P\)=\(P\)-value.
\(^b\)btwt=birth weight; Effect of birth weight modified by the variables parity and hypertension.
\(^c\)PDA=patent ductus arteriosus.
\(^d\)Hypertension = maternal hypertensive disease of pregnancy.
Table 34: Logistic regression results –Final Model External Validation Sample*

External Model

N=32,770

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.02</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.0014</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>African American</td>
<td>0.29</td>
<td>1.35</td>
<td>1.23, 1.47</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaching full enteral feeding by 5 d of age</td>
<td>-0.80</td>
<td>0.45</td>
<td>0.31, 0.65</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA* No treatment</td>
<td>0.38</td>
<td>1.44</td>
<td>1.23, 1.72</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDA Indomethacin only</td>
<td>0.07</td>
<td>1.05</td>
<td>0.96, 1.20</td>
<td>0.21</td>
</tr>
<tr>
<td>PDA Surgery</td>
<td>0.40</td>
<td>1.45</td>
<td>1.29, 1.74</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>NO PDA reported</td>
<td>REF</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>0.42</td>
<td>------</td>
<td></td>
<td>0.028</td>
</tr>
<tr>
<td>Hypertension x birth weight</td>
<td>-.001</td>
<td>------</td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hypertension = yes, decreasing weight</td>
<td>------</td>
<td>1.25</td>
<td>1.21, 1.30</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hypertension = no, decreasing weight</td>
<td>------</td>
<td>1.15</td>
<td>1.13, 1.17</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

AUC 0.668 (95% CI 0.66, 0.68)

*β = beta coefficient from logistic regression, OR = odds ratio, 95% CI = 95% confidence interval, P=p-value.

*bPDA = Patent ductus arteriosus.
Table 35: Logistic regression results infants 401-999 g—Internal Sample and External Validation Sample

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th></th>
<th></th>
<th></th>
<th>External Model</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=1595</td>
<td></td>
<td></td>
<td></td>
<td>N = 14,283</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.76</td>
<td>------------</td>
<td>------------</td>
<td>0.0005</td>
<td>-1.45</td>
<td>------------</td>
<td>------------</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.002</td>
<td>0.92&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.87, 0.98</td>
<td>0.006</td>
<td>-.0011</td>
<td>0.90&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.86, 0.93</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Surgical PDA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.613</td>
<td>1.85</td>
<td>1.09, 3.12</td>
<td>0.022</td>
<td>0.28</td>
<td>1.32</td>
<td>1.13, 1.55</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Parity of child &gt; 1</td>
<td>0.459</td>
<td>1.58</td>
<td>1.11, 2.27</td>
<td>0.012</td>
<td>0.02</td>
<td>1.02</td>
<td>0.91, 1.14</td>
<td>0.71</td>
</tr>
<tr>
<td>Antenatal Steroids</td>
<td>0.589</td>
<td>1.80</td>
<td>1.19, 2.74</td>
<td>0.006</td>
<td>0.17</td>
<td>1.18</td>
<td>1.06, 1.33</td>
<td>0.002</td>
</tr>
<tr>
<td>AUC</td>
<td>0.61 (95% CI 0.56, 0.66)</td>
<td></td>
<td></td>
<td></td>
<td>0.56 (95% CI 0.54, 0.57)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>β = beta coefficient from logistic regression, OR = odds ratio, 95% CI = 95% confidence interval, P=p-value.

<sup>b</sup>For every 100 gram increase in birth weight.

<sup>c</sup>PDA = patent ductus arteriosus.
<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>External Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=2221</td>
<td>N = 18,487</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.17</td>
<td>-0.56</td>
</tr>
<tr>
<td>Birth weight</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>African American</td>
<td>0.67</td>
<td>0.40</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>REF</td>
<td>REF</td>
</tr>
<tr>
<td>Reaching full enteral feeds in first 5 days of life</td>
<td>-1.06</td>
<td>-0.71</td>
</tr>
<tr>
<td>PDA&lt;sup&gt;c&lt;/sup&gt; No treatment</td>
<td>1.09</td>
<td>0.72</td>
</tr>
<tr>
<td>PDA&lt;sup&gt;c&lt;/sup&gt; Indomethacin only</td>
<td>-0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>PDA&lt;sup&gt;c&lt;/sup&gt; Surgery</td>
<td>2.01</td>
<td>0.99</td>
</tr>
<tr>
<td>No PDA reported</td>
<td>REF</td>
<td>REF</td>
</tr>
<tr>
<td>Maternal age &lt; 21y</td>
<td>0.40</td>
<td>0.27</td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>-0.72</td>
<td>-0.65</td>
</tr>
<tr>
<td>AUC</td>
<td>0.73 (95% CI 0.68, 0.78)</td>
<td>0.66 (95% CI 0.64, 0.68)</td>
</tr>
</tbody>
</table>

<sup>a</sup>β = beta coefficient from logistic regression, OR = odds ratio, 95% CI = 95% confidence interval, P=p-value.
<sup>b</sup>PDA = patent ductus arteriosus
<sup>c</sup>For every 100 gram increase in birth weight

Table 36: Logistic regression results infants 1000-1500 g; Internal Sample and External Validation Sample

---

159
Table 37: Comparison of Sensitivity and Specificity of Original and Externally Validated Models.

<table>
<thead>
<tr>
<th></th>
<th>All infants 401-1500 g&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Infants 401-999 g&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Infants 1000-1500 g&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original</td>
<td>External</td>
<td>Original</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>69 (63-75)</td>
<td>63 (61-65)</td>
<td>51 (43-59)</td>
</tr>
<tr>
<td>Specificity</td>
<td>57 (55-59)</td>
<td>51 (50-52)</td>
<td>62 (59-65)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>11 (10-12)</td>
<td>9 (8.6-9.4)</td>
<td>13 (10-17)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>96 (95-97)</td>
<td>95 (94.7-95.3)</td>
<td>92 (90-94)</td>
</tr>
<tr>
<td>AUC</td>
<td>69 (66-73)</td>
<td>66* (65-67)</td>
<td>61 (56-66)</td>
</tr>
</tbody>
</table>

After re-estimating the regression coefficients on the external sample<sup>d</sup>

<table>
<thead>
<tr>
<th></th>
<th>External</th>
<th></th>
<th>External</th>
<th></th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>401-1500 g&lt;sup&gt;d&lt;/sup&gt;</td>
<td>401-999 g&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1000-1500 g&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>77 (75-79)</td>
<td>63 (60-65)</td>
<td>70 (67-73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>48 (47-49)</td>
<td>45 (44-46)</td>
<td>51 (50-52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>10 (9.5-10.5)</td>
<td>12 (11.5-12.5)</td>
<td>6 (5.5-6.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>96.5 (96-97)</td>
<td>92 (91-93)</td>
<td>3 (92-93.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Values according to predicted probability cutoff of 6% or greater.
<sup>b</sup>Values according to predicted probability cutoff of 10% or greater.
<sup>c</sup>Values according to predicted probability cutoff of 4% or greater.
<sup>d</sup>Values based on re-estimated regression coefficients on external population sample, but with the same probability cutoffs as stated above.
Figure 20: Tree-based analysis of infants with NEC

Node 1

263
3816

< 1000 g

≥1000 g Birth weight

Node 2

109 (5)
2221

Not treated / Surgical PDA

PDA

No PDA / Indomethacin PDA

T-Node 1

27 (16)
173

Node 3

82 (4)
2048

Full Feeds

Yes

T-Node 2

3 (1)
334

No

Node 4

79(5)
1714

Caucasian

Race

African American

T-Node 3

7 (4)
173

No

Node 5

24(7)
517

Steroids

Yes

T-Node 4

3 (3)
92

T-Node 5

24 (10)
252

Hypertension

No

T-Node 6

2 (<1)
239

No

Node 6

27(8)
344

T-Node 8

5 (4)
127

Yes

Steroids

No

Yes

Node 8

43 (5)
958

Delivery

Intubation

Yes

Node 9

16 (7)
228

T-Node 9

11 (11)
101
Intermediate (child) and terminal nodes indicated by circles and boxes respectively. Inside each node, the node number is given along with the number of NEC cases (1st number) and the sample size of the node (bottom number). The number in parentheses represents the proportion of infants with NEC in each node. The proportions are rounded to the nearest whole number for simplicity. The splitting rules are specified under the circles.

Striped nodes indicate “high risk” nodes and clear nodes represent the “lower risk” nodes. Infants in a high risk node were more likely to be misclassified as having disease (false positives) and infants in lower risk nodes were more likely to be misclassified as not having disease (false negatives).
Table 38: Comparison of Performance of CART and Logistic Regression\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>CART</th>
<th>Logistic\textsuperscript{b}</th>
<th>Logistic\textsuperscript{c}</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>68 (64-72)</td>
<td>69 (66-73)</td>
<td>69 (66-73)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>80 (75-84)</td>
<td>69 (63-75)</td>
<td>49 (43-55)</td>
</tr>
<tr>
<td>Specificity</td>
<td>51 (49-53)</td>
<td>57 (55-59)</td>
<td>72.5 (71-74)</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>11 (10-12)</td>
<td>11 (10-12)</td>
<td>12 (10-14)</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>97 (92-98)</td>
<td>96 (95-97)</td>
<td>85 (94-96)</td>
</tr>
<tr>
<td>Overall correct classification rate</td>
<td>53 (51-54.5)</td>
<td>57.5 (56-59)</td>
<td>71 (70-72)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}All values represented as percentages with 95% confidence intervals.

\textsuperscript{b}Utilizing a 6% probability cutoff.

\textsuperscript{c}Utilizing an 8% probability cutoff.
Figure 21: Number of NEC events and Incidence Rate over time

*Incidence rate reported as number of cases per 1000 patient-days with 95% confidence intervals.
Figure 22: Overall Seasonal NEC Rate (1996-2004)

[Graph showing seasonal NEC rate with data points and error bars for Winter, Spring, Summer, and Autumn. The x-axis represents the seasons, and the y-axis represents cases per 1000 patient-days.]
Figure 23: Incidence of NEC for Study Years 1996-2000 and 2001-2004

P-value based on Pearson’s chi-square.
Table 39: Temporal Clusters identified within a Geographic Region

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Year</th>
<th># Cases</th>
<th>IDR</th>
<th>Scan p-value</th>
<th>Rate p-value</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1996</td>
<td>2</td>
<td>5.36</td>
<td>&lt;.0001</td>
<td>0.38</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>2</td>
<td>1996</td>
<td>2</td>
<td>4.74</td>
<td>&lt;.0001</td>
<td>0.104</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>3</td>
<td>1996</td>
<td>3</td>
<td>7.35</td>
<td>0.012</td>
<td>0.017</td>
<td>Consistent</td>
</tr>
<tr>
<td>4</td>
<td>1996</td>
<td>3</td>
<td>5.90</td>
<td>0.0127</td>
<td>0.072</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>5</td>
<td>1996</td>
<td>2</td>
<td>5.85</td>
<td>0.0127</td>
<td>0.073</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>6</td>
<td>1997</td>
<td>2</td>
<td>5.09</td>
<td>0.0127</td>
<td>0.093</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>7</td>
<td>1997</td>
<td>2</td>
<td>5.71</td>
<td>0.04</td>
<td>0.078</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>8</td>
<td>1997</td>
<td>2</td>
<td>6.94</td>
<td>&lt;.0001</td>
<td>0.054</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>9</td>
<td>1997</td>
<td>2</td>
<td>4.02</td>
<td>&lt;.0001</td>
<td>0.136</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>10</td>
<td>1998</td>
<td>2</td>
<td>4.82</td>
<td>0.0127</td>
<td>0.101</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>11</td>
<td>1998</td>
<td>3</td>
<td>7.33</td>
<td>0.003</td>
<td>0.017</td>
<td>Consistent</td>
</tr>
<tr>
<td>12</td>
<td>1998</td>
<td>3</td>
<td>7.06</td>
<td>0.003</td>
<td>0.019</td>
<td>Consistent</td>
</tr>
<tr>
<td>13</td>
<td>1998</td>
<td>3</td>
<td>8.29</td>
<td>0.003</td>
<td>0.012</td>
<td>Consistent</td>
</tr>
<tr>
<td>14</td>
<td>1999</td>
<td>2</td>
<td>5.42</td>
<td>0.0127</td>
<td>0.083</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>15</td>
<td>1999</td>
<td>2</td>
<td>4.58</td>
<td>0.0127</td>
<td>0.110</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>16</td>
<td>1999</td>
<td>2</td>
<td>6.92</td>
<td>0.0127</td>
<td>0.055</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>17</td>
<td>1999</td>
<td>5</td>
<td>10.50</td>
<td>0.0007</td>
<td>0.002</td>
<td>Consistent</td>
</tr>
<tr>
<td>18</td>
<td>1999</td>
<td>2</td>
<td>6.82</td>
<td>0.003</td>
<td>0.112</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>19</td>
<td>1999</td>
<td>2</td>
<td>4.67</td>
<td>&lt;.0001</td>
<td>0.108</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>20</td>
<td>1999</td>
<td>2</td>
<td>5.41</td>
<td>&lt;.0001</td>
<td>0.084</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>21</td>
<td>2000</td>
<td>2</td>
<td>4.75</td>
<td>0.0127</td>
<td>0.104</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>22</td>
<td>2000</td>
<td>2</td>
<td>5.19</td>
<td>0.0127</td>
<td>0.090</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>23</td>
<td>2000</td>
<td>3</td>
<td>5.83</td>
<td>0.0127</td>
<td>0.074</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>24</td>
<td>2000</td>
<td>3</td>
<td>10.21</td>
<td>0.0121</td>
<td>0.007</td>
<td>Consistent</td>
</tr>
<tr>
<td>25</td>
<td>2000</td>
<td>5</td>
<td>7.69</td>
<td>0.0121</td>
<td>0.015</td>
<td>Consistent</td>
</tr>
<tr>
<td>26</td>
<td>2000</td>
<td>3</td>
<td>4.40</td>
<td>0.042</td>
<td>0.116</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>27</td>
<td>2000</td>
<td>2</td>
<td>4.88</td>
<td>&lt;.0001</td>
<td>0.099</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>28</td>
<td>2000</td>
<td>2</td>
<td>5.25</td>
<td>0.0127</td>
<td>0.088</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>Cluster ID</td>
<td>Year</td>
<td># Cases</td>
<td>IDR</td>
<td>Scan p-value</td>
<td>Rate p-value</td>
<td>Cluster</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>--------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>29</td>
<td>2001</td>
<td>3</td>
<td>7.56</td>
<td>&lt;.0001</td>
<td>0.015</td>
<td>Consistent</td>
</tr>
<tr>
<td>30b</td>
<td>2001</td>
<td>4</td>
<td>7.73</td>
<td>0.023</td>
<td>0.014</td>
<td>Consistent</td>
</tr>
<tr>
<td>31</td>
<td>2001</td>
<td>2</td>
<td>4.59</td>
<td>0.042</td>
<td>0.109</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>32b</td>
<td>2001</td>
<td>4</td>
<td>6.76</td>
<td>0.0127</td>
<td>0.02</td>
<td>Consistent</td>
</tr>
<tr>
<td>33</td>
<td>2001</td>
<td>3</td>
<td>7.73</td>
<td>0.003</td>
<td>0.015</td>
<td>Consistent</td>
</tr>
<tr>
<td>34b</td>
<td>2001</td>
<td>4</td>
<td>6.38</td>
<td>0.012</td>
<td>0.009</td>
<td>Consistent</td>
</tr>
<tr>
<td>35</td>
<td>2002</td>
<td>2</td>
<td>5.28</td>
<td>&lt;.0001</td>
<td>0.086</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>36</td>
<td>2002</td>
<td>2</td>
<td>4.68</td>
<td>0.042</td>
<td>0.106</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>37</td>
<td>2002</td>
<td>3</td>
<td>7.11</td>
<td>0.029</td>
<td>0.018</td>
<td>Consistent</td>
</tr>
<tr>
<td>38b</td>
<td>2002</td>
<td>3</td>
<td>5.14</td>
<td>0.042</td>
<td>0.090</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>39</td>
<td>2003</td>
<td>2</td>
<td>3.83</td>
<td>&lt;.0001</td>
<td>0.146</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>40</td>
<td>2003</td>
<td>2</td>
<td>4.62</td>
<td>&lt;.0001</td>
<td>0.11</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>41</td>
<td>2003</td>
<td>4</td>
<td>7.14</td>
<td>0.0007</td>
<td>0.007</td>
<td>Consistent</td>
</tr>
<tr>
<td>42</td>
<td>2003</td>
<td>2</td>
<td>5.13</td>
<td>0.013</td>
<td>0.091</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>43</td>
<td>2003</td>
<td>2</td>
<td>5.52</td>
<td>&lt;.0001</td>
<td>0.083</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>44</td>
<td>2004</td>
<td>5</td>
<td>8.74</td>
<td>&lt;.0001</td>
<td>0.001</td>
<td>Consistent</td>
</tr>
<tr>
<td>45c</td>
<td>2004</td>
<td>3</td>
<td>5.79</td>
<td>0.092</td>
<td>0.030</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>46</td>
<td>2004</td>
<td>3</td>
<td>6.42</td>
<td>&lt;.0001</td>
<td>0.023</td>
<td>Consistent</td>
</tr>
<tr>
<td>47</td>
<td>2004</td>
<td>2</td>
<td>5.70</td>
<td>0.013</td>
<td>0.076</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>48</td>
<td>2004</td>
<td>2</td>
<td>4.67</td>
<td>0.042</td>
<td>0.106</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>49</td>
<td>2004</td>
<td>3</td>
<td>7.53</td>
<td>&lt;.0001</td>
<td>0.015</td>
<td>Consistent</td>
</tr>
</tbody>
</table>

*aFisher’s exact p-values based on comparison of observed incidence density rates with expected incidence density rates (IDR).

*bClusters reported as 2 overlapping groupings of NEC events. Example Cluster 32: 4 cases of 2 overlapping clusters beginning on 4/27 and ending on 5/5. IDR reported above is the 7-day rate which includes cluster 5/5. See Figure 29.

*cCluster 45 was marginally significant (p=.092) with Test 1 (Scan) and statistically significant with Test 2 (Rate) indicating an inconsistent cluster.
Figure 24: Number of Consistent and Inconsistent Clusters by Season
The proportions of cases each year that occurred within defined clusters are reported here as the black bars. No statistically significant secular trend with an increase in the proportion of cases that were occurring in clusters over years was seen. The numbers at the top of the figure represent the number of cases of NEC reported each year. The p-value represents the Cochrane-Armitage test for trend over the years with regards to the proportion of cases that were within a consistent cluster.
Figure 26: Number of Clusters and Case Size by Season

- Spring: 19 cases
- Summer: 17 cases
- Autumn: 13 cases
- Winter: 12 cases
Figure 27: Temporal Cluster ID 45: An example of the usefulness of applying 2 methods for cluster identification

NEC events circled represent 2 separate potential clusters. The first series begins on 2/2 and ends on 2/9 consisting of 5 cases. The second series begins on 2/18 and ends on 2/23 consisting of 3 cases. The probability of the second series is as follows: \( P(R \geq 3 \mid 30, 8, 7) = 0.092 \): this is the probability that 3 cases would occur in the most recent 7 days given that 8 cases total have occurred in the previous 30 days. The incidence rate, however, for this series of cases is significantly above the expected rate \( (p=0.03) \), indicating that this is an inconsistent cluster.
<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Year</th>
<th># Cases</th>
<th>IDR</th>
<th>Scan p-value</th>
<th>Rate p-value</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1996</td>
<td>2</td>
<td>9.30</td>
<td>&lt;.0001</td>
<td>0.032</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>2</td>
<td>1996</td>
<td>2</td>
<td>10.31</td>
<td>&lt;.0001</td>
<td>0.027</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>3</td>
<td>1996</td>
<td>3</td>
<td>9.95</td>
<td>&lt;.0001</td>
<td>0.028</td>
<td>Consistent</td>
</tr>
<tr>
<td>4</td>
<td>1996</td>
<td>2</td>
<td>11.83</td>
<td>&lt;.0001</td>
<td>0.021</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>5</td>
<td>1997</td>
<td>2</td>
<td>9.09</td>
<td>&lt;.0001</td>
<td>0.034</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>6</td>
<td>1997</td>
<td>2</td>
<td>16.95</td>
<td>&lt;.0001</td>
<td>0.011</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>7</td>
<td>1998</td>
<td>3</td>
<td>14.15</td>
<td>0.003</td>
<td>0.003</td>
<td>Consistent</td>
</tr>
<tr>
<td>8</td>
<td>1998</td>
<td>3</td>
<td>15.63</td>
<td>0.003</td>
<td>0.002</td>
<td>Consistent</td>
</tr>
<tr>
<td>9</td>
<td>1998</td>
<td>3</td>
<td>14.42</td>
<td>0.003</td>
<td>0.003</td>
<td>Consistent</td>
</tr>
<tr>
<td>10</td>
<td>1999</td>
<td>2</td>
<td>13.61</td>
<td>&lt;.0001</td>
<td>0.016</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>11</td>
<td>1999</td>
<td>2</td>
<td>11.24</td>
<td>0.003</td>
<td>0.023</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>12</td>
<td>1999</td>
<td>2</td>
<td>9.13</td>
<td>0.003</td>
<td>0.033</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>13</td>
<td>1999</td>
<td>3</td>
<td>8.29</td>
<td>0.003</td>
<td>0.012</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>14</td>
<td>1999</td>
<td>2</td>
<td>16.95</td>
<td>&lt;.0001</td>
<td>0.011</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>15</td>
<td>1999</td>
<td>2</td>
<td>8.81</td>
<td>&lt;.0001</td>
<td>0.036</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>16</td>
<td>1999</td>
<td>2</td>
<td>12.20</td>
<td>&lt;.0001</td>
<td>0.020</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>17</td>
<td>1999</td>
<td>2</td>
<td>13.25</td>
<td>&lt;.0001</td>
<td>0.017</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>18</td>
<td>2000</td>
<td>2</td>
<td>9.57</td>
<td>&lt;.0001</td>
<td>0.031</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>19</td>
<td>2000</td>
<td>3</td>
<td>17.14</td>
<td>0.003</td>
<td>0.002</td>
<td>Consistent</td>
</tr>
<tr>
<td>20</td>
<td>2000</td>
<td>3</td>
<td>26.6</td>
<td>&lt;.0001</td>
<td>0.0004</td>
<td>Consistent</td>
</tr>
<tr>
<td>21</td>
<td>2000</td>
<td>2</td>
<td>9.17</td>
<td>0.0127</td>
<td>0.033</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>22</td>
<td>2000</td>
<td>2</td>
<td>10.20</td>
<td>&lt;.0001</td>
<td>0.027</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>23</td>
<td>2000</td>
<td>2</td>
<td>8.51</td>
<td>0.012</td>
<td>0.038</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>24</td>
<td>2001</td>
<td>2</td>
<td>10.10</td>
<td>&lt;.0001</td>
<td>0.028</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>25c</td>
<td>2001</td>
<td>4</td>
<td>15.00</td>
<td>0.029</td>
<td>0.003</td>
<td>Consistent</td>
</tr>
<tr>
<td>26c</td>
<td>2001</td>
<td>3</td>
<td>11.30</td>
<td>0.0127</td>
<td>0.023</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>27</td>
<td>2001</td>
<td>3</td>
<td>13.64</td>
<td>0.003</td>
<td>0.003</td>
<td>Consistent</td>
</tr>
<tr>
<td>28</td>
<td>2001</td>
<td>3</td>
<td>19.35</td>
<td>&lt;.0001</td>
<td>0.001</td>
<td>Consistent</td>
</tr>
<tr>
<td>29</td>
<td>2001</td>
<td>3</td>
<td>17.24</td>
<td>&lt;.0001</td>
<td>0.002</td>
<td>Consistent</td>
</tr>
<tr>
<td>Cluster ID</td>
<td>Year</td>
<td># Cases</td>
<td>IDR(^a)</td>
<td>Scan p-value</td>
<td>Rate p-value(^b)</td>
<td>Cluster</td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>---------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>30</td>
<td>2002</td>
<td>2</td>
<td>10.15</td>
<td>&lt;.0001</td>
<td>0.022</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>31</td>
<td>2002</td>
<td>2</td>
<td>7.19</td>
<td>0.042</td>
<td>0.051</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>32</td>
<td>2002</td>
<td>3</td>
<td>21.74</td>
<td>&lt;.0001</td>
<td>0.001</td>
<td>Consistent</td>
</tr>
<tr>
<td>33</td>
<td>2003</td>
<td>2</td>
<td>9.48</td>
<td>&lt;.0001</td>
<td>0.03</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>34</td>
<td>2003</td>
<td>2</td>
<td>11.90</td>
<td>&lt;.0001</td>
<td>0.02</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>35</td>
<td>2003</td>
<td>2</td>
<td>6.06</td>
<td>0.0127</td>
<td>0.069</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>36</td>
<td>2003</td>
<td>2</td>
<td>13.51</td>
<td>&lt;.0001</td>
<td>0.016</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>37</td>
<td>2004</td>
<td>2</td>
<td>12.05</td>
<td>&lt;.0001</td>
<td>0.02</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>38</td>
<td>2004</td>
<td>3</td>
<td>8.24</td>
<td>&lt;.0001</td>
<td>0.012</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>39</td>
<td>2004</td>
<td>2</td>
<td>11.24</td>
<td>0.042</td>
<td>0.02</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>40</td>
<td>2004</td>
<td>2</td>
<td>6.62</td>
<td>&lt;.0001</td>
<td>0.059</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>41</td>
<td>2004</td>
<td>3</td>
<td>13.27</td>
<td>&lt;.0001</td>
<td>0.003</td>
<td>Consistent</td>
</tr>
</tbody>
</table>

\(^a\)Incidence Density Rates are higher than previously reported in Table 41 due to hospital specific denominators, which are lower than overall regional denominators reported in Table 41.

\(^b\)Fisher's exact p-values based on comparison of observed incidence density rates with expected incidence density rates. A p-value <0.01 was considered significant.

\(^c\)Overlapping cluster.
Table 41: Comparison of cases by temporal cluster definition\textsuperscript{a}

<table>
<thead>
<tr>
<th>Prenatal/maternal factors</th>
<th>Consistent N=57</th>
<th>P*</th>
<th>Inconsistent N=62</th>
<th>P**</th>
<th>Sporadic N=83</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>26 (22-32)</td>
<td>0.03</td>
<td>28 (22-32)</td>
<td>0.028</td>
<td>24 (20-32)</td>
</tr>
<tr>
<td>Marital status – single</td>
<td>29 (51)</td>
<td>0.14</td>
<td>29 (47)</td>
<td>0.05</td>
<td>52 (63)</td>
</tr>
<tr>
<td>Delivery Method C-section</td>
<td>38 (67)</td>
<td>0.36</td>
<td>36 (58)</td>
<td>0.91</td>
<td>49 (58)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>12 (21)</td>
<td>0.41</td>
<td>15 (24)</td>
<td>0.20</td>
<td>13 (16)</td>
</tr>
<tr>
<td>Antepartum hemorrhage</td>
<td>10 (18)</td>
<td>0.21</td>
<td>12 (19)</td>
<td>0.31</td>
<td>22 (27)</td>
</tr>
<tr>
<td>Prolonged rupture of membranes</td>
<td>10 (18)</td>
<td>0.67</td>
<td>18 (29)</td>
<td>0.23</td>
<td>17 (20)</td>
</tr>
<tr>
<td>Maternal antibiotic exposure at admission</td>
<td>36 (63)</td>
<td>0.95</td>
<td>44 (71)</td>
<td>0.29</td>
<td>52 (63)</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>47 (82)</td>
<td>0.80</td>
<td>56 (90)</td>
<td>0.11</td>
<td>67 (81)</td>
</tr>
<tr>
<td>Gravida: First Pregnancy</td>
<td>35 (61)</td>
<td>0.99</td>
<td>45 (73)</td>
<td>0.16</td>
<td>51 (61)</td>
</tr>
<tr>
<td>Parity (2 or greater)</td>
<td>43 (75)</td>
<td>0.24</td>
<td>47 (76)</td>
<td>0.21</td>
<td>55 (66)</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>20 (35)</td>
<td>0.64</td>
<td>16 (26)</td>
<td>0.47</td>
<td>26 (31)</td>
</tr>
<tr>
<td>Perinatal/Neonatal factors</td>
<td>Consistent (N=57)</td>
<td>P*</td>
<td>Inconsistent (N=62)</td>
<td>P**</td>
<td>Sporadic (N=83)</td>
</tr>
<tr>
<td>---------------------------------------------------------</td>
<td>-------------------</td>
<td>----</td>
<td>---------------------</td>
<td>-----</td>
<td>-----------------</td>
</tr>
<tr>
<td>Age of NEC – median (IQR)</td>
<td>12 (8-27)</td>
<td>0.58</td>
<td>14 (8-29)</td>
<td>0.87</td>
<td>16 (9-26)</td>
</tr>
<tr>
<td>Gender – Male</td>
<td>32 (56)</td>
<td>0.82</td>
<td>30 (48)</td>
<td>0.49</td>
<td>45 (54)</td>
</tr>
<tr>
<td>Race – Black</td>
<td>26 (46)</td>
<td>0.87</td>
<td>25 (40)</td>
<td>0.42</td>
<td>39 (47)</td>
</tr>
<tr>
<td>Birth weight in grams</td>
<td>920 (714-1090)</td>
<td>0.94</td>
<td>865 (730-1090)</td>
<td>0.98</td>
<td>900 (680-1110)</td>
</tr>
<tr>
<td>Gestational age in weeks</td>
<td>27 (25-28)</td>
<td>0.25</td>
<td>27 (25-28)</td>
<td>0.26</td>
<td>27 (25-29)</td>
</tr>
<tr>
<td>SGA</td>
<td>9 (16)</td>
<td>0.72</td>
<td>10 (16)</td>
<td>0.76</td>
<td>15 (18)</td>
</tr>
<tr>
<td>5-minute Apgar Score &lt; 7</td>
<td>23 (40)</td>
<td>0.34</td>
<td>15 (24)</td>
<td>0.27</td>
<td>27 (33)</td>
</tr>
<tr>
<td>Mechanical ventilation &gt;2 days in 1st wk</td>
<td>25 (44)</td>
<td>0.82</td>
<td>32 (52)</td>
<td>0.49</td>
<td>38 (46)</td>
</tr>
<tr>
<td>Early Infant antibiotic exposure</td>
<td>34 (60)</td>
<td>0.14</td>
<td>29 (47)</td>
<td>0.98</td>
<td>39 (47)</td>
</tr>
<tr>
<td>Indomethacin, any indication</td>
<td>19 (33)</td>
<td>0.96</td>
<td>22 (35)</td>
<td>0.83</td>
<td>28 (34)</td>
</tr>
<tr>
<td>Patent ductus arterosus (PDA) untreated</td>
<td>9 (16)</td>
<td>0.41</td>
<td>7 (11)</td>
<td>0.26</td>
<td>11 (13)</td>
</tr>
<tr>
<td>PDA indomethacin treated</td>
<td>14 (25)</td>
<td>0.92</td>
<td>13 (21)</td>
<td>0.54</td>
<td>21 (25)</td>
</tr>
<tr>
<td>Surgical PDA</td>
<td>5 (9)</td>
<td>0.74</td>
<td>10 (16)</td>
<td>0.08</td>
<td>5 (6)</td>
</tr>
<tr>
<td></td>
<td>Consistent</td>
<td></td>
<td>Inconsistent</td>
<td></td>
<td>Sporadic</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>------------</td>
<td>--</td>
<td>--------------</td>
<td>--</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>N=57</td>
<td></td>
<td>N=62</td>
<td></td>
<td>N=83</td>
</tr>
<tr>
<td>Surfactant use</td>
<td>31 (54)</td>
<td>0.79</td>
<td>32 (52)</td>
<td>0.55</td>
<td>47 (57)</td>
</tr>
<tr>
<td>Delivery room resuscitation – intubation needed</td>
<td>31 (54)</td>
<td>0.12</td>
<td>32 (52)</td>
<td>0.17</td>
<td>34 (41)</td>
</tr>
<tr>
<td>Required O₂ at 6 hrs thru 24 hrs of age</td>
<td>39 (68)</td>
<td>0.91</td>
<td>44 (71)</td>
<td>0.65</td>
<td>56 (67)</td>
</tr>
<tr>
<td>Showed resp. distress within 24 hrs of age</td>
<td>55 (96)</td>
<td>0.56</td>
<td>58 (94)</td>
<td>0.16</td>
<td>82 (99)</td>
</tr>
<tr>
<td>Need for resp. support at 24 hrs</td>
<td>48 (84)</td>
<td>0.23</td>
<td>43 (69)</td>
<td>0.38</td>
<td>63 (76)</td>
</tr>
<tr>
<td>Abnormal chest x-ray at 24hrs</td>
<td>51 (89)</td>
<td>0.043</td>
<td>49 (79)</td>
<td>0.66</td>
<td>63 (76)</td>
</tr>
<tr>
<td>Received 1ˢᵗ enteral feed ≤ 3 days of age</td>
<td>34 (60)</td>
<td>0.31</td>
<td>39 (63)</td>
<td>0.67</td>
<td>55 (66)</td>
</tr>
<tr>
<td>Achieved full feeds during NICU stay</td>
<td>43 (75)</td>
<td>0.26</td>
<td>50 (81)</td>
<td>0.70</td>
<td>69 (83)</td>
</tr>
<tr>
<td>Mean days to reach full enteral feeds</td>
<td>16 (10-29)</td>
<td>0.60</td>
<td>24.50 (13-37)</td>
<td>0.015</td>
<td>14 (9-27)</td>
</tr>
<tr>
<td>NEC-associated outcomes</td>
<td>Consistent N=57</td>
<td>P*</td>
<td>Inconsistent N=62</td>
<td>P**</td>
<td>Sporadic N=83</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-----------------</td>
<td>----</td>
<td>------------------</td>
<td>-----</td>
<td>---------------</td>
</tr>
<tr>
<td>Length of Stay</td>
<td>81 (30-110)</td>
<td>0.55</td>
<td>70 (41-111)</td>
<td>0.29</td>
<td>61 (25-104)</td>
</tr>
<tr>
<td>Late onset sepsis</td>
<td>27 (47)</td>
<td>0.18</td>
<td>29 (47)</td>
<td>0.20</td>
<td>30 (36)</td>
</tr>
<tr>
<td>Death</td>
<td>23 (40)</td>
<td>0.61</td>
<td>17 (27)</td>
<td>0.27</td>
<td>30 (36)</td>
</tr>
<tr>
<td>Surgical NEC</td>
<td>35 (61)</td>
<td>0.78</td>
<td>36 (58)</td>
<td>0.91</td>
<td>49 (59)</td>
</tr>
<tr>
<td>Spontaneous perforation</td>
<td>6 (11)</td>
<td>0.77</td>
<td>11 (24)</td>
<td>0.11</td>
<td>9 (13)</td>
</tr>
<tr>
<td>Chronic Lung Disease (36 wk)</td>
<td>13 (31)</td>
<td>0.65</td>
<td>10 (28)</td>
<td>0.46</td>
<td>16 (36)</td>
</tr>
<tr>
<td>Intracranial hemorrhage (3 or 4)</td>
<td>23 (40)</td>
<td>0.61</td>
<td>17 (27)</td>
<td>0.27</td>
<td>30 (36)</td>
</tr>
</tbody>
</table>

*aContinuous variable reported as median values (IQR = interquartile range). Categorical variables reported as N (%).

b Consistent NEC vs. Sporadic NEC comparison.

c Inconsistent NEC vs. Sporadic NEC comparison.
Figure 28: Number of cases of illnesses in the community for year 2004

The dots represent the number of NEC events recorded each week (left axis). The dashed line represents the number of pertussis events (left axis). The dotted line represents the number of RSV events and the dash-dotted line represents the number of flu events. Pertussis, RSV, and flu events were according to the number of laboratory-confirmed cases captured at one center, but considered fairly representative of occurrences within the community. The arrows represent times of consistent NEC clusters.
NEC events circled represent one cluster of overlapping 7-day windows. The index case occurred on 9/17 and the final case of this cluster occurred on 9/27. The probability is stated as the following for the first 7-day window: $P(R \geq 2 | 30, 3, 7) = 0.0127$; and for the overlapping 7-day window: $P(R \geq 3 | 30, 5, 7) = 0.0121$. The incidence rates were significantly above the expected rate, indicating the occurrence of a consistent cluster.
Literature Cited


Appendix 1: Institutional Review Board Approval Letters

DATE: September 12, 2003

TO: Ed Donovan, M.D.
Division, CPRC
Cincinnati Children’s Hospital Medical Center
Winslow Bldg., RM. R-7545

FROM: Irwin Light, M.D., Chairman
Institutional Review Board
Cincinnati Children’s Hospital Medical Center
MLC #5020

RE: PREDICTION MODELS FOR NEONATAL NECROTIZING ENTEROCOLITIS
(CHMC #03-9-7X)

The purpose of this study is to develop a mathematical model that can be used to predict a premature infant’s risk of developing necrotizing enterocolitis (NEC).

The study involves a retrospective review of existing demographic and clinical information that has been entered in the Cincinnati portion of the NICHD Neonatal Research Network Registry. This registry has previously been approved by the Cincinnati Children’s Hospital Medical Center Institutional Review Board. Data will be obtained from VLBW infants who have been admitted to the NICUs at University Hospital, Children’s Hospital and Good Samaritan Hospital.

No information defined as Protected Health Information (PHI) will be included in the study data set.

Approval is granted by the Cincinnati Children’s Hospital Medical Center Institutional Review Board for the retrospective review of existing medical information noted above. Informed consent/authorization is not required for the use of de-identified data.

IL/dj
Appendix 2: Statement of Responsibility for Completion of this Dissertation

The completion of the work required for this dissertation involved the cooperation and contribution of many individuals and groups. This statement will outline the specific tasks performed by myself and those by others on this study.

The purpose of the Generic Database Study is to provide a registry of baseline and outcome data for very low birth weight infants, based on data collected in a uniform manner from neonatal intensive care units (NICUs) at institutions participating in the NICHD Neonatal Research Network. Each center is listed in the Acknowledgements.

The Steering Committee, comprised of the Principal Investigators from each of the Clinical Centers and the Data Coordinating Center, the NICHD Program Official, and the Chairman of the Steering Committee is responsible for identifying topics for network studies, designing study protocols, monitoring study implementation and recruitment. The Generic Database Subcommittee is responsible for the design of the generic data forms and for monitoring the conduct of the study. The Network Coordinator at each site is responsible for collecting and entering the data into the Neonatal Research Network data base. Although personal communication has been limited to members of the team at the Cincinnati Center, the work of this multi-center team has been extensive as it relates to the GDB study coordination and management, and I gratefully acknowledge their support and assistance. The Research Triangle Institute (RTI) is in charge of data management and conducting analyses of data that involves multiple sites. Members of RTI work extremely closely with the GDB Subcommittee, as well as with investigators interested in conducted data analysis of GDB data.
I conducted the literature review and designed the current study, including power calculations and selection of study cohort to be included. I completed all analyses for this dissertation. For the external validation of the models that I created, I wrote all of the SAS programs and Lisa Wrage from the Research Triangle Institute ran my programs on the NICHD NRN data collected as part of the GDB study. At several points during analysis, I consulted with Dr. Mekbib Altaye, as well as members of my dissertation committee on the direction of analyses and interpretation of data. Dr. Paul Succop’s close participation in these analyses was invaluable. The production of this written dissertation was entire mine, with the help of astute comments and suggestions from my dissertation committee.
Appendix 3: Published Abstracts

Title: Using Prediction Models to Enhance Informed Consent for Prevention Trials

J Meinzen-Derr, MPH1, E Donovan, MD1 and NICHD Neonatal Research Network.
'Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States.

Presented at the Pediatric Academic Societies' Annual Meeting, April 28-May 2, 2006, San Francisco, CA.

Background: Estimating the probability of disease and treatment-associated harm could aid in the selection of individuals who would benefit most from participation in trials of preventive interventions.

Objective: To develop and use multivariable models of very low birth weight (VLBW) necrotizing enterocolitis (NEC) risk to enhance the informed consent process.

Design/Methods: Using clinical characteristics of infants enrolled in the NRN VLBW registry 1991-2004, logistic regression was used to estimate likelihood of NEC among 3816 infants from a single center. The model was validated using 39,940 infants from 22 centers. Model parameters were used in hypothetical scenarios for requesting consent to participate in trials of preventive interventions with varying risk of intervention harm. Uncertain intervention risk was assumed high.

Results: A model of optimal sensitivity/specificity contained 6 variables: birth weight (odd ratio [OR] 0.85 for each 100g increase), race (Black OR 1.2), surgical PDA (OR 1.4), maternal hypertension (OR 0.7), prolonged rupture of membranes (OR 1.2), achieving full feeds in the first 5 days (OR 0.44). In a population of VLBW infants with NEC prevalence of 7% [95% CI 6.8, 7.4], this model could help inform the consent process for studies of preventive interventions. If patients at >20% risk of NEC were approached for consent, it is estimated that 85% [95% CI 78, 92] of those randomized to intervention would not develop NEC and thus potentially be exposed unnecessarily to intervention-associated risks. Uncertain intervention risk was assumed high.

Conclusions: Due to difficulty in predicting which premature infants will develop NEC, evaluation of preventive interventions is not efficient. Any potential risk associated with the intervention is thus experienced by many infants at negligible risk. By using validated models to assist in the selection of infants for trials, the number of low risk infants exposed during trials can be reduced. Partly funded by the Gerber Foundation.

<table>
<thead>
<tr>
<th>NEC probability</th>
<th>No trial benefit (%) (false pos)</th>
<th>Missed opportunity (%) (false neg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None used</td>
<td>93</td>
<td>0</td>
</tr>
<tr>
<td>0.10</td>
<td>88</td>
<td>6</td>
</tr>
<tr>
<td>0.20</td>
<td>85</td>
<td>7</td>
</tr>
</tbody>
</table>