Intercenter Differences in Bronchopulmonary Dysplasia or Death Among Very Low Birth Weight Infants

WHAT'S KNOWN ON THIS SUBJECT: Bronchopulmonary dysplasia is a major cause of morbidity in premature infants. Outcomes of most neonatal disorders, including bronchopulmonary dysplasia, vary in incidence across neonatal centers even after adjustment for demographic and antenatal characteristics. The factors responsible for such variation have not been systematically evaluated.

WHAT THIS STUDY ADDS: Bronchopulmonary dysplasia/death rates were found to demonstrate a moderate clustering effect according to center, as did clinical variables associated with the outcome. Persistent center differences after clustered variables were corrected indicated the presence of as-yet unmeasured center variables.

AUTHORS: Namasivayam Ambalavanan, MD,^a Michele Walsh, MD, MS,^b Georgiy Bobashev, PhD,^c Abhik Das, PhD,^c Burton Levine, MS,^c Waldemar A. Carlo, MD,^a and Rosemary D. Higgins, MD,^d for the NICHD Neonatal Research Network

^aDepartment of Pediatrics, University of Alabama at Birmingham, Birmingham, Alabama; ^bDepartment of Pediatrics, Case Western Reserve University, Cleveland, Ohio; ^cRTI International, Research Triangle Park, North Carolina; and ^dEunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network, Bethesda, Maryland

KEY WORDS

logistic models, infant, premature, predictive value of tests, clustering

ABBREVIATIONS

BPD—bronchopulmonary dysplasia
NRN—Neonatal Research Network
GDB—generic database
ALR—alternating logistic regression
OR—odds ratio
PWOR—pairwise odds ratio
AUC—area under the curve

www.pediatrics.org/cgi/doi/10.1542/peds.2010-0648

doi:10.1542/peds.2010-0648

Accepted for publication Sep 29, 2010

Address correspondence to Namasivayam Ambalavanan, MD, 176F Suite 9380, 619 South 20th St, University of Alabama at Birmingham, Birmingham, AL 35249. E-mail: ambal@uab.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2011 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

Funded by the National Institutes of Health (NIH).

abstract

OBJECTIVES: To determine (1) the magnitude of clustering of bronchopulmonary dysplasia (36 weeks) or death (the outcome) across centers of the Eunice Kennedy Shriver National Institute of Child and Human Development National Research Network, (2) the infant-level variables associated with the outcome and estimate their clustering, and (3) the center-specific practices associated with the differences and build predictive models.

METHODS: Data on neonates with a birth weight of <1250 g from the cluster-randomized benchmarking trial were used to determine the magnitude of clustering of the outcome according to alternating logistic regression by using pairwise odds ratio and predictive modeling. Clinical variables associated with the outcome were identified by using multivariate analysis. The magnitude of clustering was then evaluated after correction for infant-level variables. Predictive models were developed by using center-specific and infant-level variables for data from 2001–2004 and projected to 2006.

RESULTS: In 2001–2004, clustering of bronchopulmonary dysplasia/death was significant (pairwise odds ratio: 1.3; $P < .001$) and increased in 2006 (pairwise odds ratio: 1.6; overall incidence: 52%; range across centers: 32%–74%); center rates were relatively stable over time. Variables that varied according to center and were associated with increased risk of outcome included lower body temperature at NICU admission, use of prophylactic indomethacin, specific drug therapy on day 1, and lack of endotracheal intubation. Center differences remained significant even after correction for clustered variables.

CONCLUSION: Bronchopulmonary dysplasia/death rates demonstrated moderate clustering according to center. Clinical variables associated with the outcome were also clustered. Center differences after correction of clustered variables indicate presence of as-yet unmeasured center variables. *Pediatrics* 2011;127:e106–e116

Bronchopulmonary dysplasia (BPD) is a major cause of morbidity in premature infants.¹ Most neonatal outcomes such as BPD, retinopathy of prematurity, necrotizing enterocolitis, and neurodevelopmental impairment vary in incidence across neonatal centers even after adjustment for demographics and antenatal characteristics.^{2–4} The magnitude of these center variations in outcome is larger than the effect sizes of many clinically proven interventions such as surfactant,⁵ continuous positive airway pressure,⁶ or vitamin A.⁷ It is therefore important to determine the characteristics of center variation and identify whether these variations are amenable to intervention.

A benchmarking and multimodal quality improvement trial by the Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network (NRN) showed that adoption of “potentially better practices” by the intervention centers changed practices but did not improve rates of survival free of BPD.⁸ This benchmark study, which was the impetus for our study, revealed that rates of BPD were clustered within centers, ie, some centers had higher and some lower than the average rate, and this variation was larger than that expected by chance.

We hypothesized that a multivariable combination of individual characteristics and center-specific practices leads to the differences in center-specific rates of survival free of BPD at 36 weeks’ postmenstrual age (the outcome of interest). The study objectives were to (1) determine the magnitude of clustering of BPD/death across the neonatal centers, (2) determine the infant-level variables associated with the outcomes and estimate their clustering across the centers, and (3) build predictive models and determine the center-specific practices associated with the outcomes.

PATIENTS AND METHODS

This study used data collected as part of the benchmarking cluster-randomized trial⁸ performed in 17 centers of the NRN, which are academic urban NICUs with >500 admissions of infants per year (the majority of infants were admitted to the NICU of the hospital where they were born). In the benchmarking study, 3 centers with the highest rates of survival free of BPD were identified as the benchmark centers. The remaining 14 centers were randomly assigned to control ($n = 7$) or intervention groups ($n = 7$). The study population for these analyses included neonates with a birth weight of 401 to 1250 g who were born alive at these 14 centers between March 1, 2001, and April 30, 2004 ($n = 2836$).⁸ In addition, infants with a similar birth weight range ($n = 2793$) who were born in 2006 in NRN centers were analyzed as a validation group. The NRN centers in 2001–2004 and in 2006 were not the same, because membership in the NRN subject to competitive renewal. During this period, 4 centers exited the NRN, and 4 new centers entered and were evaluated in 2006. Because the benchmarking trial had ended in 2006, only variables collected routinely as part of the NRN generic database (GDB) were available for analysis in the 2006 data set.

In this study, BPD in infants at 36 weeks’ postmenstrual age was defined by using a validated physiologic definition.⁹ The outcome of BPD/death was used rather than BPD alone because death is a competing outcome for BPD (infants who die early cannot develop BPD). BPD severity was not evaluated because it was not clearly or consistently defined, and would also require the use of a multilevel outcome that would be difficult to interpret.

Variables associated with the magnitude of immaturity (prematurity related), delivery room practices, lung

injury, lung edema, lung repair, steroid administration, nutrition, infection, mechanical ventilation, and clinical practices (Table 1) were analyzed. All variables that were evaluated were determined at the level of the infant (individual), except for the center variable (which was the same for all infants born at that center).

Analysis was performed in 3 stages:

1. In stage 1 we estimated the magnitude of clustering of BPD/death outcome in centers, adjusted and unadjusted for basic demographic factors. Variables that were associated with survival free of BPD and also varied according to center were identified, as described subsequently.
2. The variables identified in stage 1 were used in a regression model to determine if they accounted for observed variation among centers.
3. A reduced predictive model was developed that included variables that were most strongly associated with BPD/death.

For stage 1, we used alternating logistic regression (ALR) with pairwise odds ratios (PWORs) as a measure of clustering. ALR is a form of logistic regression^{10,11} in which the clustering parameter is calculated in an iterative fashion together with the estimates of the regression parameters. PWORs range from 0 to infinity. A PWOR value of 1 corresponded to no clustering. Generally, a PWOR of <1.2 is considered a low magnitude of clustering, between 1.2 and 1.9 as moderate, and >1.9 as high. For variable selection we used logistic regression with a center indicator as a fixed-effect categorical variable.

For stage 2, logistic regression analysis for BPD/death was done with center indicators as fixed effects; birth weight, gender, and race as control indicators; and variables considered rel-

TABLE 1 Variables Analyzed in the Study

Variable	Category	Remarks
1 Maternal age ^a	Prematurity related	Used in the models
2 Maternal pregnancy history, gravida	Prematurity related	Used but later removed from the final model during stepwise selection
3 Maternal marital status	Prematurity related	Nonsignificant interaction with center
4 Maternal medical insurance	Prematurity related	Nonsignificant interaction with center
5 At least 1 prenatal visit	Prematurity related	Nonsignificant interaction with center
6 Maternal insulin-dependent diabetes	Prematurity related	Used but later removed from the final model during stepwise selection
7 Maternal hypertension ^a	Prematurity related	Used in the models
8 Antepartum hemorrhage	Prematurity related	Used but later removed from the final model during stepwise selection
9 Chorioamnionitis	Infection	Nonsignificant interaction with center
10 Rupture of membranes	Prematurity related	Not enough information contained in the variable
11 Antenatal steroids	Prematurity related/steroids	Not enough information contained in the variable
12 Antenatal antibiotics	Prematurity related	Used but later removed from the final model during stepwise selection
13 Final mode of delivery	Prematurity related	Used but later removed from the final model during stepwise selection
14 Birth weight ^a	Prematurity related	Used in the models
15 Gestational age	Prematurity related	Used but later removed from the final model during stepwise selection
16 Gender ^a	Prematurity related	Used in the models
17 Racial category	Prematurity related	Used but later removed from the final model during stepwise selection
18 Apgar score at 1 min	Prematurity related	Used but later removed from the final model during stepwise selection
19 Apgar score at 5 min	Prematurity related	Used but later removed from the final model during stepwise selection
21 Birth resuscitation, oxygen	Delivery room	Only 1.5% of subjects had oxygen
22 Birth resuscitation, bag and mask	Delivery room	Nonsignificant interaction with center
23 Birth resuscitation, CPAP	Delivery room	Nonsignificant interaction with center
24 Birth resuscitation, intubation ^a	Delivery room	Used in the models (note: same as variable 52, "intubation in delivery room")
25 Birth resuscitation, chest compression	Delivery room	Nonsignificant interaction with center
26 Birth resuscitation, epinephrine	Delivery room	Not enough information contained in the variable 0.63% of subjects were "Y"
27 Respiratory distress syndrome <24 h	Injury	Nonsignificant interaction with center
28 O ₂ /invasive mechanical ventilation >6 h in first 24 h	Injury	Nonsignificant interaction with center
29 Surfactant doses	Injury	Used but later removed from the final model during stepwise selection
30 Pneumothorax	Injury	Used but later removed from the final model during stepwise selection
31 Pulmonary hemorrhage	Injury	Nonsignificant interaction with center
32 Postnatal steroids, day 7 ^a	Repair/Steroids	Used in the models (day 7)
33 Patent ductus arteriosus	Edema	Nonsignificant interaction with center
34 Patent ductus arteriosus, Indocin	Edema	Nonsignificant interaction with center
35 Patent ductus arteriosus, ibuprofen	Edema	Not enough information contained in the variable 0% of subjects were "Y"
36 Patent ductus arteriosus, surgery	Edema	Nonsignificant interaction with center
37 Indomethacin prophylaxis ^a	Edema	Used in the models
38 Early onset sepsis (<72 h)	Infection	Nonsignificant interaction with center
39 Antibiotics >5 d	Infection	Used in the models
40 Late-onset culture-negative sepsis; antibiotic therapy >5 d	Infection	Nonsignificant interaction with center
41 Late-onset sepsis episodes	Infection	Nonsignificant interaction with center
42 Maximum % weight loss in first 10 d (birth weight — lowest weight)	Edema/fluids	Nonsignificant interaction with center
43 Date birth weight regained	Edema/fluids	Nonsignificant interaction with center
44 Parenteral alimentation total days	Nutrition/fluids	Nonsignificant interaction with center
45 Enteral feeds, age started (date of birth — date of starting)	Nutrition	Nonsignificant interaction with center
46 Enteral feeds, age achieving 120 mL/kg per d	Nutrition	Not enough information contained in the variable
47 Proven NEC	Infection	Nonsignificant interaction with center

TABLE 1 Continued

Variable	Category	Remarks
48 Spontaneous gastrointestinal perforation without NEC	Infection	Not enough information contained in the variable
49 Day 3, 7; days on HFV, CV, nasal synchronized intermittent mandatory ventilation, CPAP, supplemental O ₂ ^a	Mechanical ventilation	Only day 3 and day 7 variables were considered. Used in the models: no. of days on CV at day 7, on HFV at day 7, on supplemental O ₂ at day 7. Used but later removed from the final model during stepwise selection: CPAP day 3, 7; CV day 3; HFV day 3; supplemental O ₂ day 3. Nonsignificant interaction with center for all variables.
50 Place of delivery	Delivery room	Nonsignificant interaction with center (not enough information)
51 Circumstances of delivery	Delivery room	Nonsignificant interaction with center (not enough information)
52 Intubation in delivery room ^a	Delivery room	Used in the models (note: same as variable 24, "birth resuscitation intubation")
53 CPAP in delivery room	Delivery room	Used but later removed from the final model during stepwise selection
54 Surfactant in delivery room	Delivery room	Used but later removed from the final model during stepwise selection
55 Temperature on admission ^a	Prematurity related	Used in the models
56 Blood pressure on admission	Prematurity related	Used but later removed from the final model during stepwise selection
57 Respiratory settings on admission	Mechanical ventilation	Used but later removed from the final model during stepwise selection
58 First blood gas analysis results	Mechanical ventilation	Used but later removed from the final model during stepwise selection
59 Ventilation changes before blood gas	Other clinical practice	Used but later removed from the final model during stepwise selection
60 Daily respiratory support, days 1, 3, 7	Mechanical ventilation	Used but later removed from the final model during stepwise selection
61 Daily respirator support (if yes): measured mean airway pressure, days 1, 3, 7	Mechanical ventilation	Nonsignificant interaction with center
62 Daily supplemental O ₂ (if yes): FiO ₂ (if no supplemental O ₂ , entered as 0.21, "birth resuscitation, oxygen"), days 1, 3, 7 ^a	Mechanical ventilation	Used in the models: days on supplemental O ₂ by day 7 Used but later removed from the final model during stepwise selection: supplemental O ₂ day 3
63 Daily blood gases, days 1, 3, 7	Mechanical ventilation	Nonsignificant interaction with center
68 Parenteral and enteral intake data: days 1, 3, 7, 14, total intake (mL/kg per day)	Nutrition	Nonsignificant interaction with center, mainly because of much missing data.
69 Parenteral and enteral intake data, days 1, 3, 7, 14: parenteral intake (mL/kg per day)	Nutrition	Nonsignificant interaction with center, mainly because of much missing data.
70 Parenteral and enteral intake data, days 1, 3, 7, 14: enteral intake (mL/kg per day)	Nutrition	Nonsignificant interaction with center, mainly because of much missing data.
71 Parenteral and enteral intake data, days 1, 3, 7, 14: parenteral intake: amino acids (g/kg per day)	Nutrition	Nonsignificant interaction with center, mainly because of much missing data.
72 Parenteral and enteral intake data, days 1, 3, 7, 14: caloric density	Nutrition	Nonsignificant interaction with center, mainly because of much missing data.
73 Score for Neonatal Acute Physiology	Prematurity related	Nonsignificant interaction with center, mainly because of much missing data.
74 Neonatal therapeutic intervention score form on day 1, all items: I (respiratory), II (cardiovascular), III (drug therapy), IV (monitoring), V (metabolic/nutrition), VI (transfusion), VII (procedures), VIII (vascular access) ^a	Prematurity related; other clinical care practices	Used in the models
75 Training of most senior member of neonatal team	Delivery room/other clinical care practices	Nonsignificant interaction with center
76 Training of person who actually delivered baby	Delivery room/other clinical care practices	Nonsignificant interaction with center
77 Resuscitation with epinephrine or sodium bicarbonate or saline in delivery room	Delivery room/other clinical care practices	Nonsignificant interaction with center
78 Urine output at day 1 (mL/kg per day) ^a	Prematurity related	Used in the model

Variables that did not have a significant effect on BPD/death and had no significant interaction with the center variable were initially screened out. We used a stepwise procedure to select a parsimonious set of variables with minimal collinearity. Variables that were present in at least 1 of the final models are noted as "used in the models." Variables 50 through 78 were available only for the 2001–2004 data set as part of the benchmarking study. HFV indicates high-frequency ventilation; CPAP, continuous positive airway pressure; "Y", yes; NEC, necrotizing enterocolitis; CV, conventional ventilation.

^a Variables that were used in any of the presented models.

evant to BPD/death. The center variable was used as a categorical indicator; therefore the centers were compared simultaneously. We added an interaction term between the variable of interest (the candidate variable) and the center variable to examine the significance of variable interactions associated with BPD/death. The interaction term determines whether the association of the candidate variable with BPD/death is moderated according to the center at which the particular practice or risk factor is located. We used only variables defined earlier than or at day 7 to reduce confounding with the outcome (eg, ventilatory requirements on day 14 are strongly associated with BPD/death and would be confounders). With these identified variables, ALR was used to calculate PWOR.

For stage 3, we used a stepwise selection process to develop a reduced (parsimonious) model for BPD/death using the variables that were defined within the first 7 days and with at least 1 of the following: significant association (adjusted for birth weight, gender, race) with the outcome, or significant interaction with the center variable when it was added to the model. A stepwise selection algorithm was then run with a conservative $P < .05$ threshold for both “enter” and “stay,” which used these selected variables and no center indicator. We also estimated the statistical significance for the center variable by explicitly adding it to the model. For model validation we repeated a stepwise variable selection algorithm based on GDB data to develop parsimonious models for the 2006 data and compared the model with the 1 for 2001–2004. Both 2001–2004 and 2006 data sets contain the same standard GDB variables, and these variables were used for the validation of the predictive model. However, additional variables were available in 2001–2004

TABLE 2 Characteristics of the Infants

Variable	2001–2004	2006
Birth weight, <i>n</i> ; mean $\pm$ SD, g	2836; 910 $\pm$ 206	2794; 874 $\pm$ 230
Gestational age, <i>n</i> ; median (25th–75th centile), wk	2836; 27.1 (25–29)	2794; 26.7 (25–29)
Male (1 = male), <i>n</i> (%)	2836 (51.2)	2794 (49.8)
Black not Hispanic, <i>n</i> (%)	2836 (42.9)	2794 (38.8)
BPD, <i>n</i> (%)	2812 (34)	2085 (39.3)
Death, <i>n</i> (%)	2828 (14.4)	2777 (25.1)
BPD/death, <i>n</i> (%)	2836 (35.6)	2794 (52.2)

from the benchmarking study, and these variables could not be evaluated in the validation set. All analyses were performed with SAS 9.2 (SAS Institute, Inc, Cary, NC).

RESULTS

Patient characteristics and sample sizes for the 2001–2004 and 2006 cohorts are presented in Table 2. The sample sizes were similar between the 2001–2004 and 2006 cohorts. The incidence of BPD was similar in the 2 time periods, but the combined outcome of BPD/death was higher in 2006 because of increased mortality in the different group of NRN centers.

Stage 1: Estimation of the Magnitude of Clustering and the Selection of Variables for Inclusion in the Models

In 2001–2004 clustering of BPD/death within the centers was moderate (PWOR: 1.3; $P < .001$; overall incidence: 35%; range across centers: 16%–54%). Clustering was greater in 2006 (PWOR: 1.6; $P < .001$; overall incidence: 52%; range: 32%–74%), which indicated that prevalence of BPD/death varied significantly across centers. Even after we accounted for birth weight, gender, and race, significant ($P < .0001$) center variation for BPD/death persisted for both time periods (Table 3). Using center as a fixed-effect categorical variable, we examined specific between-center variations (Table 3). If birth rate, gender, or race variation between centers were responsible for the difference, then we expected the

adjusted OR to be closer to 1. However, we observed the opposite effect for some centers, for which the adjusted OR was higher than unadjusted ORs. This result suggests that variation in birth weight, race, or gender was not the reason for the outcome variation between the centers. The rates of BPD/death in centers that were in the NRN at both time periods (2001–2004 and 2006) were relatively stable, ie, we did not observe significant variation over time within centers.

The variables that had a statistically significant bivariate association with the outcome (BPD/death), as well as an interaction between the variable and the center (at $P \leq .05$), are shown in Table 4. These variables were candidates for additional analysis because they indicate both risk of BPD/death and variation according to center.

Stage 2: Regression Model

Logistic regression analysis was done with center indicators as fixed effects; birth weight, gender, and race as control indicators; and the variables of interest from Table 4. After addition of all candidate variables the center variable was still significant at $P < .0001$, and the magnitude of clustering was not markedly reduced (PWOR: 1.42 for 2001–2004 and 1.33 for 2006).

Stage 3: Predictive Model

The reduced models for 2001–2004 data are presented in Table 5. When the center variable was added, “mother’s age,” “admission temperature,”

TABLE 3 Prevalence of BPD/Death, Raw ORs, and ORs Adjusted for Birth Weight, Gender, and Race

Center	<i>n</i>	Mean Birth Weight, g	BPD/Death, %	Raw OR	Adjusted OR	95% CI	<i>P</i>
For 2001–2004							
A	229	885	39.7	1.18	1.11	0.82–1.51	.49
B	266	946	30.1	0.77	0.85	0.63–1.13	.26
C	223	910	53.8	2.09	2.86	2.09–3.92	<.0001
D	133	900	31.6	0.83	0.98	0.65–1.47	.92
E	189	936	41.8	1.29	1.40	1.01–1.93	.05
F	158	904	50.0	1.79	1.822	1.29–2.58	.001
G	167	888	42.5	1.33	1.01	0.71–1.43	.96
H	116	861	38.8	1.14	1.10	0.74–1.66	.63
I	287	828	42.2	1.31	1.39	1.06–1.82	.02
J	180	896	16.1	0.34	0.25	0.17–0.39	<.0001
K	340	783	32.9	0.88	0.75	0.58–0.97	.03
L	219	903	32.0	0.84	1.10	0.80–1.51	.55
X	300	936	22.0	0.51	0.46	0.34–0.63	<.0001
For 2006							
M	123	852	50.4	0.93	0.97	0.63–1.49	.88
N	151	857	51.0	0.95	0.92	0.62–1.35	.67
O	188	857	55.9	1.16	1.28	0.89–1.83	.18
P	199	881	48.7	0.87	0.97	0.68–1.38	.86
Q	285	904	47.7	0.84	0.99	0.74–1.31	.92
R	187	861	66.3	1.80	2.23	1.56–3.19	<.0001
S	152	858	53.3	1.05	1.06	0.73–1.55	.77
T	158	836	50.6	0.94	0.68	0.46–0.99	.04
U	129	913	47.3	0.82	0.83	0.55–1.26	.38
V	254	880	31.5	0.42	0.32	0.24–0.44	<.0001
W	267	835	63.3	1.58	1.54	1.14–2.09	.01
MM	119	882	31.9	0.43	0.36	0.23–0.55	<.0001
MN	56	894	46.4	0.79	0.69	0.38–1.26	.23
MO	32	909	43.8	0.71	0.71	0.31–1.63	.42
MP	63	865	42.9	0.69	0.45	0.25–0.81	.008
MQ	109	857	55.0	1.12	0.96	0.62–1.50	.87
MR	57	897	68.4	1.99	2.56	1.38–4.78	.003
MS	152	927	74.3	2.66	4.33	2.86–6.56	<.0001
MT	75	943	74.7	2.70	5.23	2.93–9.33	<.0001
Z	38	834	34.2	0.48	0.32	0.15–0.68	.003

One center with <30 infants was removed. The reference category for the ORs is the mean center-level prevalence.

TABLE 4 Variables That Showed Both Significant Association With BPD/Death and Significant Interaction With the Center Indicators in the 2001–2004 Data

Variables	Model Without Interaction		Model With Interaction	
	OR	Main-Effect <i>P</i>	Main-Effect <i>P</i>	Variable-Center Interaction <i>P</i>
Total No. of surfactant doses, 1 vs >1	0.35	<.0001	<.001	<.05
NICU admission: blood pressure, diastolic	0.98	<.0001	<.001	<.01
NICU admission: blood pressure, systolic	0.97	<.0001	<.001	<.05
First ABG, PacO ₂	1.01	<.0001	<.001	<.05
NICU surfactant given	3.56	<.0001	<.001	<.05
Mean airway pressure, day 3	0.995	<.0001	<.001	<.05
Respiratory support, day 3 at 6:00	0.89	<.0001	<.001	<.01
Respiratory support, day 3 at 12:00	0.91	<.0001	<.001	<.01
Respiratory support, day 3 at 18:00	0.92	<.0001	<.001	<.05
Urine output first 12 h	0.80	<.0001	<.001	<.01
Indomethacin prophylaxis	3.66	<.0001	<.001	<.05
Monitoring intensity score, day 1	1.76	<.0001	<.001	<.05
Endotracheal intubation	1.59	<.0001	<.001	<.05
No. of late-onset culture-negative infection	1.79	<.0001	<.001	<.05
Days on conventional ventilation by day 3	1.71	<.0001	<.001	<.01
Days on supplemental O ₂ by day 7	1.60	<.0001	<.001	<.001

Interaction with the center variable suggested that these variables are likely to be responsible for clustering of outcomes within the centers. ABG indicates arterial blood gas.

“drug therapy, day 1,” “prophylactic indomethacin” and “endotracheal intubation” lost significance, which confirmed a strong association between the center and these variables (Table 5). The drug-therapy variable was a weighted sum of multiple medications (antimicrobial agents, diuretics, steroids, anticonvulsants, aminophylline, metabolic acidosis therapy, and other medications) to indicate the aggressiveness of pharmacotherapy.

Validation

The area under the curve (AUC) of the model with the use of benchmarking data from 2001–2004 was high (0.89), which indicated good predictive accuracy. The reduced or parsimonious model with the use of variables available in both 2001–2004 and 2006 data sets produced an AUC of 0.84 on the 2001–2004 data set and an AUC of 0.84 for the 2006 data set, which indicated good model validity. Although the predictive power was strong, a model with the single variable of birth weight produced an AUC of 0.78, which indicated that birth weight was the strongest predictor of BPD/death.

Validation based on multivariate parsimonious models indicated that many variables showed consistent cross-year significance (Table 6). At the same time, it was not possible to validate the consistent significance of certain variables (eg, prophylactic indomethacin, endotracheal intubation) that were significant in earlier analysis performed with the benchmarking data (Table 5) because these variables were available only in the benchmarking data set.

Overlap of certain variables was noted in the multivariate models, but other variables were unique to either the 2001–2004 model or 2006 model (Table 6), which indicated that the change in composition of centers during these 2 time periods, as well as temporal

TABLE 5 Variables in the Reduced Predictive Model for 2001–2004 Data, Which Included Variables From the GDB and Benchmarking Studies

Variable	OR	P
Birth weight (per gram)	.997	<.0001
Male	1.38	<.05
Mother's age, y ^a		<.05
<17 vs >35	2.43	—
18–20 vs >35	0.95	—
21–28 vs >35	1.03	—
29–34 vs >35	0.76	—
Maternal hypertension/preeclampsia/eclampsia	0.69	<.05
NICU admission temperature ^a	0.76	<.001
Drug therapy, day 1 ^a	0.85	<.01
Indomethacin prophylaxis ^a	1.92	<.001
Endotracheal intubation ^a	0.63	<.01
Postnatal steroids by day 7	2.95	<.0001
Days on conventional ventilation by day 7	1.16	<.001
Days on HFV by day 7	1.32	<.0001
Days on supplemental oxygen by day 7	1.12	<.05

^a Variables that lost significance when the center was added, which indicated a strong association between the center and these practices.

TABLE 6 Variables Common to 2001–2004 and 2006 Data That Were Used in Validation and Show Cross-Year Significance

Variable Name	Variables Significant in the Multivariate Model	
	2001–2004 Data Set	2006 Data Set
Birth weight, g ^a	Passed at .05	Passed at .05
Days on supplemental O ₂ by day 7 ^a	Passed at .05	Passed at .05
No. of late-onset culture-negative infections ^a	Passed at .05	Passed at .05
Days on high-frequency ventilation by day 7 ^a	Passed at .05	Passed at .05
Days on conventional ventilation by day 7 ^a	Passed at .05	Passed at .05
Pregnancy complications: hypertension ^a	Passed at .05	Passed at .05
Male (1 = male) ^a	Passed at .05	Passed at .05
No. of days on continuous positive airway pressure by day 7	Passed at .05	Rejected at .2
Black non-Hispanic	Rejected at .2	Passed at .05
Mother's age	Rejected at .05	Rejected at .2
Gravida	Rejected at .05	Rejected at .2
Apgar score at 5 min	Passed at .05	Rejected at .2
Pregnancy complications: diabetes	Rejected at .05	Rejected at .2
Days on high-frequency ventilation by day 3	Passed at .05	Rejected at .2
Days on conventional ventilation by day 3	Passed at .05	Rejected at .2
Days on CPAP by day 3	Rejected at .05	Rejected at .2
Patent ductus arteriosus	Rejected at .2	Passed at .05
Days on supplemental O ₂ by day 3	Rejected at .2	Passed at .05
Infant received antibiotics for ≥5 d	Rejected at .2	Passed at .05
Indomethacin prophylaxis	Rejected at .2	Rejected at .2
No. of surfactant doses	Rejected at .2	Rejected at .2

A stepwise algorithm was used to select the parsimonious model for each of the years.

^a Variables that show significance ("passed at .05") at .05 level in both 2001–2004 and 2006.

changes in clinical practice, may have led to changes in the predictive model.

DISCUSSION

BPD/death rates demonstrated strong clustering effects according to center that persisted over time. Some clinical variables associated with BPD/death

were also clustered. The persistence of center differences after correction of clustered variables indicates the presence of as-yet unmeasured center variables that contribute to the center differences.

The strengths of this study included the multicenter nature of the study,

which involved 14 academic NICUs from different states, the prospective cohort evaluation, the use of a validated physiologic definition of BPD in combination with a "hard" outcome such as mortality, the use of novel statistical methods, and the validation on a more recent cohort of infants. The existence of intercenter variations in important neonatal outcomes has been recognized for many years, and the magnitude of the differences in these outcomes is often far greater than the effect of any therapeutic intervention,^{3,4,12} but the factors responsible for such variation have not been systematically evaluated to date. Systematic evaluation of the factors underlying center variation is also lacking in the pediatric and adult intensive-care literature.

A limitation of this study was that limited data on details of clinical practices were obtained. Most studies tend to collect information on patient characteristics and outcomes but relatively little on clinical practices (eg, indications for mechanical ventilation, oxygen saturation targets) that may have an impact on patient outcomes. The order in which particular treatments and interventions are performed is important in regard to determination of possible causality and the best decision-making strategies. Although our results may indicate associations between outcomes and risk factors, the causal relationships are difficult to establish. For example, birth weight and gestational age were correlated but birth weight was a slightly better predictor of survival and thus stepwise selection algorithms selected birth weight to stay in the models. Birth weight measurements are an estimate of gestational age as well as intrauterine growth, and hence may be more strongly associated with outcome than gestational age alone. However, decisions about prophylactic sur-

factant administration are usually made by using gestational age. Such a relationship may have been lost in the analysis.

Clustering of outcomes means that the outcome in some centers was significantly higher or lower than in the others and that the difference was unlikely to be attributable to chance. This situation is different from what occurs with clustering in variables when several variables vary similarly between the individuals in a center. Classic statistical analysis relies on an assumption of independent and identically distributed observations from a sample. Often this assumption is violated by the fact that observations from the same hospital or even region are more similar to each other than the observations from other hospitals or regions and thus cannot be considered independent. A number of factors may lead to such clustering, for example, similar clinical practices, patient characteristics, or unit culture.

Determination of the magnitude of clustering of BPD/death may be done in a simple manner by an analysis of variance and estimation of the variation within and between centers and tests of whether the variation in outcomes between centers could be attributable to chance. Although simple to perform, such an analysis would not allow identification of critical covariates associated with the noted differences. Thus, a more sophisticated statistical method was employed. We first analyzed between-center clustering using PWORs, and we then identified individual and center-level variables that were associated with the outcome and clustering, and we finally partially validated the models by repeating some of the analysis on a new data set. In regression analysis of continuous outcomes, clustering is measured by intraclass correlation, which is not an efficient measure for binary (yes/no)

outcomes because variance depends on the prevalence. PWOR is a more appropriate measure of clustering of binary outcomes.^{10,11} PWOR determines the likelihood that 2 infants with the same outcome will be found together in 1 center compared with random chance (ie, PWOR indicates how much the odds of an outcome increases within a cluster compared with completely random allocation). ALR is a regression method that enables investigators to simultaneously examine the impact of the independent variables, account for clustering in the calculation of SEs, and quantify the amount of clustering, with PWOR as a measure of clustering.

Most variables investigated in this study were defined at the individual (infant) level, so when a variable shows variation between the centers it is demonstrating the variation of these individual characteristics across centers. Although the exact information about center-level practices is not available, individual-level procedures may indicate center-level practices. The goal of this analysis was not to point out specific centers and suggest changes of practices, but rather to raise awareness that certain practices (or combinations of practices) should be analyzed in more detail at the center level rather than at the individual level.

Recent nonrandomized clinical quality-improvement studies in which investigators aimed at improving the performance of ICUs for specific outcomes such as nosocomial infection^{13,14} or BPD/death¹⁴ have led to some success in reducing adverse outcomes, although the cluster-randomized benchmarking trial⁸ did not demonstrate such benefit. In general, these studies have implemented “potentially better practices” developed through analysis of the processes of patient care, literature review, and clinical site visits.¹⁴

However, these studies had the limitations that (1) not all beneficial practices may be identified as “potentially better practice,” perhaps because of lack of evidence, because large, expensive clinical trials may be required to confirm small effects of changes in clinical practice; (2) some practices may either not improve outcome or may possibly adversely affect outcome at certain centers, because of center-specific variation in other practices that may interact with the practice being implemented (eg, prophylactic indomethacin reduces severe intraventricular hemorrhage,¹⁵ but in combination with early postnatal steroids may increase risk of gastrointestinal perforation¹⁶); and (3) many common practices in neonatal intensive care show much variation in implementation. In general, variables contributing to intercenter differences (after correction for illness severity and patient characteristics) can be broadly classified as: (1) known measured variables, which are known to be associated with outcomes and are commonly measured (eg, antenatal steroids, use of continuous positive airway pressure, and severity of intraventricular bleed); (2) known unmeasured variables, which are known to be associated with outcomes but are not commonly measured, either to limit the number of variables for analysis or to limit the cost and labor involved (eg, tidal volumes, blood gas variables, blood cytokine concentrations, and magnitude of histologic chorioamnionitis); (3) unknown variables, (eg, unknown markers of illness severity, clinician expertise, experience of nurses, staffing patterns, medicolegal environment, and moral and ethical beliefs of caregivers and parents). The contribution of the known measured, known unmeasured, and unknown variables to BPD/death may vary according to postnatal age and other modifying variables.¹⁷ For example, oxygen satura-

tion targets and the compliance with these targets influence mortality and other outcomes (eg, retinopathy of prematurity), as demonstrated in recent studies (eg, the recent SUPPORT [Surfactant Positive Airway Pressure and Pulse Oximetry Randomized Trial] trial).¹⁸ In general, most centers have oxygen saturation targets of 85% to 95%, but exact targets were not pre-specified in the benchmarking study, and oxygen saturation targets in infants show much intracenter as well as intercenter variation, as well as variation in the same infant over time. The variables that were identified as predictors of BPD/death and which lost significance when the center variable was added, which indicated a strong association between the center and these variables, were maternal age, admission temperature, drug therapy on day 1, prophylactic indomethacin, and endotracheal intubation. Higher admission temperature, use of more medication, maternal age range of 29 to 34 years, and endotracheal intubation were associated with less BPD/death, whereas use of prophylactic indomethacin was associated with BPD/death. It is not clear if the variables associated with less BPD/death indicate a greater willingness on the part of clinicians to resuscitate aggressively. We need to use caution in the interpretation of multivariate analysis because the final model will depend on available data and specific type of analysis. Many variables are strongly correlated and the presence of 1 variable could mask the effect of another variable that might have less statistical significance but more practical importance. It must also be remembered that use of an intervention does not necessarily indicate “need” for that intervention, and that association of a variable with a bad outcome may indicate either that the variable is a marker of outcome

(and not a predisposing factor), or that it is a response of the neonate or the clinician to a different factor associated with the bad outcome.

CONCLUSIONS

We know a great deal about risk factors for BPD and for death in very pre-term infants. However, the magnitude of the center variation in BPD/death is large, and center differences persist after correction of clinical variables associated with BPD/death. Additional research into the clustered and unmeasured variables, perhaps by thoughtful and careful prospective data collection and intervention trials, may yield insights into the pathogenesis and prevention of BPD/death. Similar analyses of intercenter variation in outcomes in critically ill populations are required to identify the underlying factors responsible for variation in outcome and determine processes to attenuate these differences and improve overall outcomes.

ACKNOWLEDGMENTS

The National Institutes of Health and the Eunice Kennedy Shriver National Institute of Child Health and Human Health and Development provided grant support for the NRN’s benchmarking study (recruitment March 1, 2001, to April 30, 2004) and generic database study (recruitment January to December 2006). This work was supported by grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Health and Development and the Department of Health and Human Services (U10 HD21385, U10 HD40689, U10 HD27871, U10 HD21373, U01 HD36790, U10 HD40498, U10 HD40461, U10 HD34216, U10 HD21397, U10 HD27904, U10 HD40492, U10 HD27856, U10 HD40521, U10 HD27853, U10 HD27880, U10 HD27851) and from the National Institutes of Health (GCRC M01 RR 08084, M01 RR 00125, M01 RR 00750, M01 RR

00070, M01 RR 0039–43, M01 RR 00039, and 5 M01 RR00044). Data collected at participating sites of the Eunice Kennedy Shriver National Institute of Child Health and Human Health and Development NRN were transmitted to RTI International, the data coordinating center for the network, which stored, managed and analyzed the data for this study. On behalf of the NRN, Drs Abhik Das (data coordinating center principal investigator) and Georgiy Bobashev (data coordinating center statistician) had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis.

The following investigators, in addition to those listed as authors, participated in this study (National Institutes of Health grant numbers are listed also): NRN chair: Alan Jobe, MD PhD, University of Cincinnati; Cincinnati Children’s Hospital Medical Center, University of Cincinnati Hospital and Good Samaritan Hospital (GCRC M01 RR08084, U10 HD27853): Edward F. Donovan, MD, Kurt Schibler, MD, Kathleen Bridges, MD, Barbara Alexander, RN, Cathy Grisby, BSN, CCRC, Jody Hessling, RN, Holly L. Mincey, RN, BSN, and Marcia Worley Mersmann, RN, CCRC; Duke University School of Medicine University Hospital, Alamance Regional Medical Center, and Durham Regional Hospital (GCRC M01 RR30, U10 HD40492): Ronald N. Goldberg, MD, C. Michael Cotten, MD, MHS, and Kathy Auten, MSHS; Emory University Children’s Health Care of Atlanta, Grady Memorial Hospital, and Emory Crawford Long Hospital (GCRC M01 RR39, U10 HD27851): Barbara J. Stoll, MD, Susie Buchter, MD, Ellen C. Hale, R, RN, BS, CCRC; Eunice Kennedy Shriver National Institute of Child Health and Human Development: Linda L. Wright, MD, Rosemary D. Higgins, MD, Elizabeth M. McClure, MD; Indiana University Indiana University Hospital, Methodist Hospital, Riley Hospital for

Children, and Wishard Health Services (GCRC M01 RR750, U10 HD27856): James A. Lemons, MD, Brenda B. Poin-dexter, MD, MS, William A. Engle, MD, Diana D. Appel, RN, BSN, Dianne Herron, RN, Lucy Miller, RN, BSN, CCRC, Richard Hooper, RRT; Rainbow Babies & Chil-dren's Hospital (GCRC M01 RR80, U10 HD21364): Avroy A. Fanaroff, MD, Nancy S. Newman, RN; RTI International (U01 HD36790): W. Kenneth Poole, PhD, Betty Hastings, Elizabeth McClure, MD, Qing Yao, PhD, Jeanette O'Donnell Auman, BS, Carolyn Petrie Huitema, MS, Scott E. Schaefer, MS, Kristin M. Zaterka-Baxter, RN; Stanford University Lucile Packard Children's Hospital (GCRC M01 RR70, U10 HD27880): David K. Stevenson, MD, Krisa P. Van Meurs, MD, William D. Rhine, MD, M. Bethany Ball, BS, CCRC, Carol Kibler, RN, Jeffrey R. Parker, RRT; Tufts Medical Center Floating Hospital for Children (GCRC M01 RR54, U10 HD53119): Ivan D. Frantz, III, MS, Brenda L. MacKinnon, RNC, Ellen Nylen, RN, University of Ala-bama at Birmingham Health System and Children's Hospital of Alabama (GCRC M01 RR32, U10 HD34216): Monica V. Collins, RN, BSN, MaEd, Shirley S. Cosby, RN, BSN; University of California, San Diego Medical Center and Sharp Mary Birch Hospital for Women (U10 HD40461): Neil N. Finer, MD, Maynard R. Rasmussen, MD, Paul

R. Wozniak, MD, Greg Heldt, MD, Kathy Arnell, RNC, Clarence Demetrio, RN, Chris Henderson, RCP, CRTT, Wade Rich, BSHS, RRT, Mindy Grabarczyk, BSN, Christina Joseph, RRT, Renee Bridge, RN, Jim Goodmar, RRT; Univer-sity of Iowa Children's Hospital (GCRC M01 RR59, U10 HD53109): Edward F. Bell, MD, Karen J. Johnson, RN, BSN; University of Miami Holtz Children's Hospital (GCRC M01 RR16587, U10 HD21397): Shahnaz Duara, MD, Ruth Everett-Thomas, RN, MSN; University of New Mexico Health Sciences Center (GCRC M01 RR997, U10 HD53089): Kristi L. Watterberg, MD, Conra Backstrom Lacy, RN; University of Rochester Gol-isano Children's Hospital at Strong (GCRC M01 RR44, U10 HD40521): Dale L. Phelps, MD, Robert A. Sinkin, MD, Linda Reubens, RN; University of Texas South- western Medical Center at Dallas Park- land Health & Hospital System and Children's Medical Center Dallas (GCRC M01 RR633, U10 HD40689): Abbot R. Laptook, MD, Walid A. Salhab, MD, Charles R. Rosenfeld, MD, Pablo J. Sánchez, MD, James Allen, RRT, Alicia Guzman, Gay Hensley, RN, Susie Madi- son, RN, Melissa Martin, RN, Nancy A. Miller, RN; University of Texas Health Science Center at Houston, Medical School, Children's Memorial Hermann Hospital, and Lydon Baines Johnson Gen- eral Hospital (U10 HD21373): Jon E. Ty-

son, MD, MPH, Kathleen A. Kennedy, MD, MPH, Esther G. Akpa, RN, BSN, Patty A. Cluff, RN, Claudia I. Franco, RNC, MSN, Anna E. Lis, RN, BSN, Georgia E. McDavid, RN, Nora I. Alaniz, BS, Patti Pierce Tate, RCP, University of Utah University Hospi- tal, LDS Hospital, and Primary Children's Medical Center (GCRC M01 RR64, U10 HD53124): Roger G. Faix, MD, Bradley A. Yoder, MD, Karen A. Osborne, RN, BSN, Jennifer J. Jensen, RN, BSN, Wake Forest University Baptist Medical Center, Bren- ner Children's Hospital, and Forsyth Medical Center (GCRC M01 RR7122, U10 HD40498): T. Michael O'Shea, MD, MPH, Nancy J. Peters, RN, CCRP; Wayne State University Hutzel Women's Hospital and Children's Hospital of Michigan (U10 HD21385): Seetha Shankaran, MD, Re-becca Bara, RN, BSN, Geraldine Muran, RN, BSN, S. Nadya Kazzi, MD, MPH, Kim-berly Hayes-Hart, RN, MSN, NNP-BC, Maria Batts, RRT; Women & Infants Hos- pital of Rhode Island (U10 HD27904): Wil- liam Oh, MD, Abbot R. Laptook, MD, An- gelita Hensman, BSN, RNC; Yale University Yale-New Haven Children's Hospital (GCRC M01 RR125, GCRC M01 RR6022, U10 HD27871): Richard A. Ehren- kranz, MD, Patricia Gettner, RN, Monica Konstantino, RN, BSN, JoAnn Poulsen, RN. We are indebted to our medical and nursing colleagues and the infants and their parents who agreed to take part in this study.

REFERENCES

1. Fanaroff AA, Stoll BJ, Wright LL, et al. Trends in neonatal morbidity and mortality for very low birthweight infants. *Am J Obstet Gynecol*. 2007;196(2):147.e1–e8
2. Horbar JD, McAuliffe TL, Adler SM, et al. Variability in 28-day outcomes for very low birth weight infants: an analysis of 11 neonatal intensive care units. *Pediatrics*. 1988;82(4): 554–559
3. Vohr BR, Wright LL, Dusick AM, et al. Center differences and outcomes of extremely-low-birth-weight infants. *Pediatrics*. 2004; 113(4):781–789
4. Cotten CM, Oh W, McDonald S, et al. Pro- longed hospital stay for extremely prema- ture infants: risk factors, center differ- ences, and the impact of mortality on selecting a best-performing center. *J Peri- natol*. 2005;25(10):650–655
5. Soll RF. Prophylactic natural surfactant ex- tract for preventing morbidity and mortal- ity in preterm infants. *Cochrane Database Syst Rev*. 2000;(2):CD000511
6. Ho JJ, Subramaniam P, Henderson-Smart DJ, Davis PG. Continuous distending pres- sure for respiratory distress syndrome in preterm infants. *Cochrane Database Syst Rev*. 2002;(2):CD002271
7. Darlow BA, Graham PJ. Vitamin A supplemen- tation for preventing morbidity and mortality in very low birthweight infants. *Cochrane Da- tabase Syst Rev*. 2002;(4):CD000501
8. Walsh M, Laptook A, Kazzi SN, et al. A cluster- randomized trial of benchmarking and mul- timodal quality improvement to improve rates of survival free of bronchopulmonary dysplasia for infants with birth weights of less than 1250 grams. *Pediatrics*. 2007; 119(5):876–890
9. Walsh MC, Yao Q, Gettner P, et al. Impact of a physiologic definition on bronchopulmo- nary dysplasia rates. *Pediatrics*. 2004; 114(5):1305–1311
10. Carey V, Zeger SL, Diggle P. Modelling multi- variate binary data with alternating logistic regressions. *Biometrika*. 1993;80(3): 517–526
11. Bobashev GV, Anthony JC. Use of alternating lo-

- gistic regression in studies of drug-use clustering. *Subst Use Misuse*. 2000;35(6–8):1051–1073
12. Hack M, Horbar JD, Malloy MH, Tyson JE, Wright E, Wright L. Very low birth weight outcomes of the National Institute of Child Health and Human Development Neonatal Network. *Pediatrics*. 1991;87(5):587–597
 13. Bloom BT, Craddock A, Delmore PM, et al. Reducing acquired infections in the NICU: observing and implementing meaningful differences in process between high and low acquired infection rate centers. *J Perinatol*. 2003;23(6):489–492
 14. Horbar JD, Rogowski J, Plsek PE, et al. Collaborative quality improvement for neonatal intensive care. NIC/Q Project Investigators of the Vermont Oxford Network. *Pediatrics*. 2001;107(1):14–22
 15. Schmidt B, Davis P, Moddemann D, et al. Long-term effects of indomethacin prophylaxis in extremely-low-birth-weight infants. *N Engl J Med*. 2001;344(26):1966–1972
 16. Stark AR, Carlo WA, Tyson JE, et al. Adverse effects of early dexamethasone in extremely-low-birth-weight infants. National Institute of Child Health and Human Development Neonatal Research Network. *N Engl J Med*. 2001;344(2):95–101
 17. Ambalavanan N, Carlo W. How can we prevent early death in preterm infants? *J Pediatr (Rio J)*. 2008;84(4):283–285
 18. Carlo WA, Finer NN, Walsh MC, et al. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med*. 2010;362(21):1959–1969