Predictive Model Selection for Tracheostomy or Death in Neonates with Severe Bronchopulmonary Dysplasia (sBPD)

Angel(Anqi) Zheng[1]
[1]Brown University School of Public Health, RI

Overview

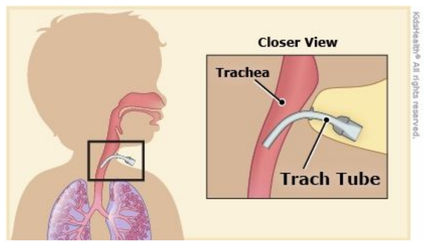
Infants with sBPD often undergo tracheostomy to facilitate their living after discharge. This project performs model and variable selection to develop a predictive model using hospital data including: data at birth-level, at 36 & 44 weeks postmenstrual age, and at discharge.

Background

- **Severe Bronchopulmonary Dysplasia (sBPD)** impacts over 10,000 infants annually
- Due to **premature**, their **lung** is under-developed structurally then lead to dependence for ventilator support
- **~75%** of infants with sBPD discharged from hospital with **tracheostomy**, a surgical procedure that creates opening on neck to allow airway
- This procedure has risks of **death** and **infection**
- Therefore a predictive model that determines **who** of the patients actually need tracheostomy and **when** to give tracheostomy is crucial
- what model fits? What variables matter?

Study Population

- **United States and Sweden**
- BPD Collaborative Registry: multi-center consortium, contains patients data across different medical centers
- 994 patients from 9 medical centers
- Dependent on ventilator at 36 weeks corrected PMA



About the Variables

Some centers have patients that are more severe ventilation support and medication status.

Birth Variables	Discharge Variables
• weight • gestational age • length and head circumference • delivery method • gender • maternal race • prenatal steroids • chorioamnionitis	• **outcome** tracheostomy death • gestational age

at 36 weeks PMA	at 44 weeks PMA
• weight • ventilation support level • fraction of inspired oxygen (FiO2) • peak inspiratory pressure • positive and exploratory pressure (PEEP) • medication for pulmonary hypertension (PH)	

Methods (Model Selection)

- **Multiple Imputation:**
  fill in values for the missing entries based on algorithm
  → a completed dataset
- **Cross Validation:**
  iteratively partitioning the dataset to test for model performance
  → robust model performance evaluation
- Each **model selection methods** is combined with multiple imputation and cross validation

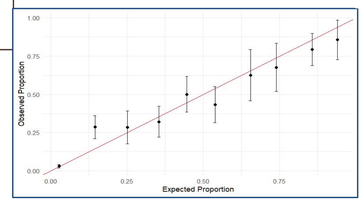
LASSO (least absolute shrinkage and selection operator) | **Best Subset** | **Forward Selection**

→ a model that minimizes the method-specific test error

Result Model

Variable	Coefficient	Variable (Cont)	Coefficient (Cont)
intercept	-5.202	**weight** (36wk)	3 × 10-4
birth length	0.047	**non-invasive VS** (36wk)	-0.177
birth head circumference	0.010	**invasive VS** (36wk)	1.788
cesarean delivery	0.652	**FiO2** (36wk)	1.508
prenatal corticosteroids	0.342	**weight** (44wk)	5.8 × 10-4
complete prenatal steroids	0.428	**non-invasive VS** (44wk)	-0.391
male	0.063	**invasive VS** (44wk)	0.905
surfactant in 72hrs	0.039	**FiO2** (44wk)	0.064
non-invasive VS * medication for PH	0.623	**PEEP** (44wk)	0.170
invasive VS * medication for PH	-1.141	**medication for PH** (44wk)	0.800
FiO2 * medication for PH	2.286		

Upon comparison of the three model selection methods, **LASSO** yields best predictive model:



Conclusion

- **LASSO** is a better model selection method than Best Subset and Forward, in this setting
- the final predictive model involves data **across timepoints** and **interactions at 44 weeks**
- some variables were excluded by the LASSO model selection process
- calibration plot above shows that the result model is relatively accurate
- coefficients for weights are small, might not be an important predictor
- more predictors are from **44 weeks** PMA, might be a more important timing than birth and 33 weeks
- **with this model, hospitals can predict whether the patient needs tracheostomy based on their respiratory and birth variables to avoid unnecessary surgical risks**