Modeling Tracheostomy Prevalence for Better Health Outcomes

Tara Amruthur

Abstract

This study utilizes a national dataset of 996 infants with severe Bronchopulmonary Dysplasia (sBPD) from 10 centers, focusing on demographics, diagnostics, and respiratory parameters at 36 and 44 weeks post-menstrual age to predict if an infant should receive a tracheostomy. Concerns arise from a disproportionate center representation, potential biases, and notable tracheostomy percentage variations among centers.

Variable selection identifies significant covariates, emphasizing prenatal corticosteroids, small for gestational age (SGA), inspired O2, medication for pulmonary hypertension (PH), and invasive positive pressure. Correlation analyses guide the exclusion of correlated metrics, and missing data patterns prompt careful imputation considerations.

Fixed effects models and mixed-effects models, acknowledging center-level variation, reveal consistent predictors in both 36 and 44-week models. Evaluation metrics favor the 44-week model, indicating superior predictive accuracy, precision, and recall. Surprisingly, the addition of a random intercept for center minimally affects model performance on validation data, suggesting limited center-level variation.

This study emphasizes the complexity of predicting tracheostomy in sBPD infants, providing direction for future predictive models for tracheostomy.

Introduction

Bronchopulmonary dysplasia (BPD), also known as chronic lung-disease, is known to cause long-term breathing problems in premature babies. Unfortunately, this can lead to poor growth and development. BPD often happens in premature babies, because their lungs are often not developed fully at the time of birth, and they need help breathing. These babies can receive breathing assistance from either a mechanical ventilator or oxygen. However, in most cases, BPD develops after a premature baby receives this breathing assistance for a period of time, as it can damage their lungs, which are already fragile (n.d.).

In recent years, there has been an increase in the survival of infants with severe BPD (sBPD), leading to significant increases in long-term ventilation due to sBPD. This has become the most common indication for tracheostomy in infants less than one year of age. There has been evidence highlighting that tracheostomy in sBPD may lead to improvements in short- and long-term respiratory and neuro-developmental outcomes. However, there are many factors that can impact whether tracheostomy is appropriate for infants, such as timing of liberation from the ventilator, odds of decannulation, rate of rehospitalization, growth, and neurodevelopment (Akangire and Manimtim 2023).

Hence, the imperative lies in discerning the suitability of an infant for tracheostomy and identifying the optimal timing for such an intervention. Our objective is to harness clinical data obtained at 36 and 44 weeks post-menstrual age (PMA) to forecast the likelihood of subsequent tracheostomy necessity or mortality before discharge.

Data

Data Collection & Properties

The study utilized a national dataset encompassing demographic, diagnostic, and respiratory parameters of infants diagnosed with sBPD. The dataset comprises of 996 distinct records, detailing information such as center identification, maternal race and ethnicity, birth metrics (weight, gestational age, length, head circumference), delivery method, prenatal corticosteroid usage, and other relevant metrics. These records pertain specifically to infants admitted to collaborative Neonatal Intensive Care Units (NICUs) and encompass known respiratory support parameters at 36 weeks and 44 weeks post-menstrual age (PMA).

The data is sourced from 10 distinct centers. Notably, a substantial proportion (64%) of the dataset is derived from a single center, while the remaining 36% is distributed among the other centers. This distribution

disparity raises the potential for bias, particularly if the characteristics of the predominantly represented center significantly differ from those of the remaining centers. This lopsided representation may impact the generalizability and potential biases in the findings derived from this analysis.

Outliers are discernible within the dataset. Employing standardized data and employing a threshold of 3.29 for comparison, it is evident that outliers exist for variables encompassing birth weight, inspiratory pressure at 36 weeks, positive end-expiratory pressure (PEEP) at 36 weeks, and inspired oxygen levels at both time points.

The distribution of tracheostomy percentages among the various centers displays notable heterogeneity. Center 1 exhibits a considerably higher incidence of tracheostomy, with 42% of infants undergoing the procedure, in contrast to Center 2, where only 10% of infants have undergone the same procedure. This disparity suggests a potential association between the specific medical center and the observed outcomes. Plausible explanations for this variability include the likelihood that certain centers function as referral facilities, consequently catering to cases of greater clinical severity, while others may not serve in this capacity, possibly receiving infants with milder conditions. Such variations underscore the necessity of integrating the variation by center into the predictive model.

Across the entirety of the dataset, the observed proportion of tracheostomy instances gives rise to apprehension. A mere 14% of the infants in the cohort have undergone tracheostomy, posing a potential impediment to the precision of any predictive model aimed at accurately forecasting tracheostomy events.

Exploratory Data Analysis

Preceding the model construction, an initial phase of variable selection was undertaken, involving an examination of correlations between covariates and variations in these covariates based on tracheostomy status. The aim was to discern which covariates merited consideration for inclusion in the model. Table 1 provides a comprehensive overview of the disparities in covariates, accompanied by the corresponding p-values derived from Chi-Squared tests. The statistical analysis reveals significant differences (at $\alpha = 0.05$) in covariates between infants who underwent tracheostomy and those who did not, encompassing variables such as birth weight, delivery method, prenatal corticosteroids, small for gestational age (SGA), weight at 36 weeks, ventilation support levels at 36 and 44 weeks, inspired oxygen at 36 and 44 weeks, peak inspiratory pressure at 36 and 44 weeks, positive and exploratory pressure (PEEP) at 36 and 44 weeks, medication for pulmonary hypertension (PH) at 36 and 44 weeks, and hospital discharge gestational age.

Furthermore, Figure 1 illustrates correlations among continuous variables, indicating a noteworthy correlation (with $\rho > 0.8$) between birth weight, birth length, and birth head circumference. To mitigate multicollinearity concerns, only birth weight was retained for analysis, deemed more informative for this investigation.

Additionally, a correlation analysis underscores the interdependence of variables collected at both the 36 and 44-week intervals. Notably, the weight at 36 weeks exhibits a high correlation ($\rho = 0.8$) with weight at 44 weeks. In response, two separate models were devised, each focusing on covariates collected at distinct time points, namely 36 weeks and 44 weeks, to enhance model reliability and interpretability.

Table 1. Differences by Tracheostomy Status

Characteristic	No, N = 850[1]	Yes, N = 146[1]	p-value[2]
Hispanic or Latino Mother	66 (8.2%)	8 (5.9%)	0.3
Birth Weight	753 (606, 950)	685 (560, 870)	0.004
Cesarean Delivery	594 (70%)	114 (79%)	0.035
Prenatal Corticosteroids	712 (86%)	123 (94%)	0.011
Complete Prenatal Steroids	524 (76%)	86 (76%)	>0.9
Maternal Chorioamnionitis	138 (17%)	22 (16%)	0.8
Male	498 (59%)	86 (59%)	>0.9
Small for Gestational Age	161 (19%)	42 (30%)	0.005
Any Surfactant	395 (81%)	66 (88%)	0.14

Weight (36 Weeks)	2,140 (1,864, 2,405)	2,030 (1,793, 2,289)	0.025
Ventilation Support Level (36 Weeks)			<0.001
Invasive Positive Pressure	168 (20%)	92 (72%)	
No Respiratory Support	111 (13%)	6 (4.7%)	
Non-Invasive Positive Pressure	560 (67%)	29 (23%)	
Inspired O2 (36 Weeks)	0.30 (0.23, 0.35)	0.45 (0.35, 0.60)	<0.001
Peak Inspiratory Pressure (36 Weeks)	0 (0, 0)	14 (2, 23)	<0.001
Positive and Exploratory Pressure (36 Weeks)	7 (6, 8)	8 (6, 9)	<0.001
Medication for Pulmonary Hypertension (36 Weeks)	41 (4.9%)	25 (20%)	<0.001
Weight (44 Weeks)	3,730 (3,260, 4,130)	3,630 (3,163, 4,015)	0.2
Ventilation Support Level (44 Weeks)			<0.001
Invasive Positive Pressure	71 (15%)	86 (77%)	
No Respiratory Support	262 (57%)	7 (6.3%)	
Non-Invasive Positive Pressure	128 (28%)	18 (16%)	
Inspired O2 (44 Weeks)	0.28 (0.25, 0.32)	0.39 (0.30, 0.58)	<0.001
Peak Inspiratory Pressure (44 Weeks)	0 (0, 0)	18 (9, 37)	<0.001
Positive and Exploratory Pressure (44 Weeks)	0 (0, 8)	8 (8, 10)	<0.001
Medication for Pulmonary Hypertension (44 Weeks)	48 (10%)	51 (46%)	<0.001
Hospital Discharge Gestational Age	45 (42, 52)	72 (55, 100)	<0.001

[1] n (%); Median (IQR)

[2] Pearson's Chi-squared test; Wilcoxon rank sum test
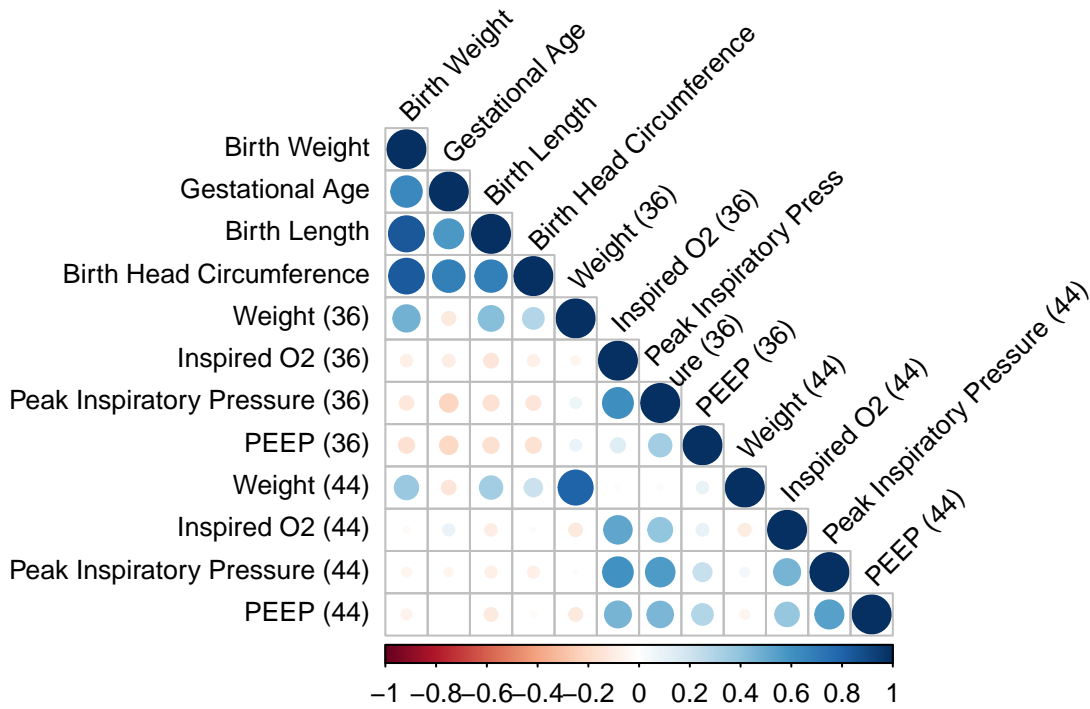

Figure 1. Correlation Plot Between all Continuous Covariates

Missing Data

The data derived from diverse medical centers manifests noteworthy disparities, with Center 4 notably lacking recorded measurements at the 44-week interval. Furthermore, a substantial proportion of centers exhibit a heightened prevalence of missing data points for measurements obtained at the 44-week interval compared to those collected at the 36-week mark. Specifically, approximately 43% of observations contain missing

values for all measurements collected at the 44-week mark. These patterns of missing data warrant thorough investigation and, if deemed appropriate, consideration of imputation techniques.

Table 3 delineates the eight columns with the highest percentage of missing data in the entire dataset. Remarkably, more than half of these columns comprise measurements collected at 44 weeks, posing potential challenges for subsequent analysis. The remaining columns involve the infant's receipt of surfactant in the first 72 hours and complete prenatal steroids. Examination of Figure 1 reveals correlations between measurements at 44 weeks and those at 36 weeks, facilitating the imputation of missing data for the former. Similarly, the relationship between the infant's receipt of complete prenatal steroids and prenatal corticosteroids allows for reliable imputation. However, the absence of variables enabling the accurate prediction of whether an infant received surfactant renders imputation unreliable in this particular context, prompting its exclusion from the imputation process and subsequent analysis.

Table 2: Columns with Largest Volume of Missing Data

Column	Percent Missing
Inspired Oxygen (44)	44.98
Peak Inspiratory Pressure (44)	44.98
Weight (44)	44.78
PEEP (44)	44.78
Any Surfactant	43.47
Ventilation Support (44)	42.57
Medication for PH (44)	42.57
Complete Prenatal Steroids	19.38

Methodology

Model Creation

For the analytsis, we opted for a training-validation split ratio of 80-20. Nevertheless, an inherent class imbalance is evident in the dataset, with less than a quarter of instances featuring infants who underwent tracheostomy. Consequently, it becomes imperative to maintain commensurate proportions of the outcome of interest within both the training and validation subsets. After partitioning the data, the training dataset encompassed 15% of infants who received a tracheostomy, while the validation dataset contained 14% of instances with infants undergoing tracheostomy. This meticulous balancing approach aims to mitigate potential biases arising from the class imbalance during subsequent model training and evaluation phases.

For model training and variable selection, the best subsets approach was used. Given a response vector $Y \in \mathbf{R}^n$, predictor matrix $X \in \mathbf{R}^{n \times p}$ and a subset size k between 0 and $\min\{n, p\}$, this approach finds the k predictors that best fit the data, in terms of the squared error. Essentially, it solves the non-convex problem:

$$\underset{\beta \in \mathbf{R}^p}{\text{minimize}} \quad ||Y - X\beta||_2^2 \quad \text{subject to} \quad ||\beta||_0 \leq k.$$

(Hastie, Tibshirani, and Tibshirani 2020).

As previously outlined, two distinct models were developed: one focused on predictors gathered at 36 weeks, and the other on predictors obtained at 44 weeks. A combined approach involving best subsets and 10-fold cross-validation was employed for variable selection in each model to enhance model generalizability, mitigate overfitting, and yield more robust estimates. This iterative procedure transpired across the five imputed datasets, with subsequent calculation of average coefficients for the selected variables. The resultant averages were then utilized as the definitive covariates in the subsequent model validation phase.

While the consideration of interaction terms was intrinsic to the model selection process, the best subsets method exhibited a propensity for favoring interaction terms over main effects. Consequently, the decision

was made to exclusively focus on main effects in order to prioritize model interpretability and coherence.

Model Validation

To assess the model's efficacy on the test dataset, a suite of metrics including precision, recall, F-score, and the Area Under the Curve (AUC) were employed. The Receiver Operating Characteristic (ROC) curve facilitated the identification of the optimal threshold for predicted probabilities to predict tracheostomy. These metrics were computed individually for each of the five imputed test datasets and subsequently averaged to yield definitive values. This comprehensive evaluation process was applied to both models that were fitted, ensuring a robust and representative assessment of model performance.

Mixed-Effects Model

The observed variation across different centers appears to be of sufficient magnitude to justify its consideration in the model. While incorporating center as a covariate or creating separate models for each center are conceivable alternatives, they pose challenges by compromising the model's generalizability. In instances where the objective is to model variable variation within a hierarchical dataset, particularly one characterized by distinct levels such as individuals (infants) and centers, a Generalized Linear Mixed Effects (GLMM) model becomes a fitting choice.

To construct the GLMM, the fixed effects models, determined through best subsets variable selection, underwent modification to encompass a random intercept for the center. This approach acknowledges and accommodates the hierarchical structure inherent in the data, ensuring a nuanced exploration of variability at both the individual and center levels. The models were trained on an amalgamated dataset, consolidating information from all five imputed test datasets. They were evaluated based on the same metrics as the fixed effect models and the Intra-Class Coefficient.

Results

Model Results

Table 3 presents the coefficients derived from the best subsets variable selection process for both fixed effects models. In the 36-week model, pivotal predictors include prenatal corticosteroids, small for gestational age (SGA), inspired oxygen, medication for pulmonary hypertension (PH), and invasive positive pressure. Conversely, the 44-week model emphasizes variables such as cesarean section, prenatal corticosteroids, SGA, PEEP, medication for PH, and ventilation support invasive positive pressure.

The consistent selection of prenatal corticosteroids and SGA in both models underscores their significance in predicting tracheostomy. Notably, medication for PH and invasive positive pressure emerge as significant predictors at both temporal points.

Analysis of Table 3 reveals a positive effect on the log odds of receiving a tracheostomy for all covariates in both models. Specifically, according to the 36-week model, infants exposed to prenatal corticosteroids are 1.44 times more likely to undergo tracheostomy, while those administered medication for PH exhibit a 1.39 times higher likelihood compared to their counterparts. Invasive positive pressure yields the largest odds ratio, signifying a substantial difference between infants receiving this respiratory support and those devoid of respiratory assistance. Despite the anticipated association between being classified as SGA and tracheostomy, SGA infants are 1.07 times more likely to undergo tracheostomy than their non-SGA counterparts.

Within the 44-week model, invasive positive pressure emerges with the highest odds ratio. Unexpectedly, a pronounced difference in the odds of tracheostomy is observed for infants administered medication for PH compared to the 36-week model. Additionally, maternal chorioamnionitis surfaces as a predictor in the 44-week model, despite its absence in the 36-week model.

Table 3: Coefficients for Fixed Effect Models: 36 Weeks (Left) and 44 Weeks (Right)

	Value		Value
Prenatal Corticosteroids	0.3628571	Cesarean Section	0.0735943
SGA	0.0685680	Prenatal Corticosteroids	1.0757771
Growth (from Birth)	-0.2446522	Maternal Chorioamnionitis	-0.1135059
Inspired O2	0.4831512	SGA	0.1688241
Medication for PH	0.3345652	Growth (from 36 Weeks)	-0.2011117
Invasive Positive Pressure	1.7550211	PEEP	0.6536269
		Medication for PH	0.9996653
		Invasive Positive Pressure	1.4257997

The addition of β_{0j}, the random intercept where j indicates center, resulted in the following two models.

$$E[Y_0] = \beta_{0j} + \beta_1 \times X_{\text{prenatal corticosteroids}} + \beta_2 \times X_{\text{SGA}} + \beta_3 \times X_{\text{growth(from birth)}} + \beta_4 \times X_{\text{inspired O2}} +$$
$$\beta_5 \times X_{\text{medication for PH}} + \beta_6 \times X_{\text{non-invasive positive pressure}} + \beta_7 \times X_{\text{invasive positive pressure}}$$

for the 36 week time points and

$$E[Y_0] = \beta_{0j} + \beta_1 \times X_{\text{cesarean section}} + \beta_2 \times X_{\text{prenatal corticosteroids}} + \beta_3 \times X_{\text{maternal chorioamnionitis}} + \beta_4 \times X_{\text{SGA}} +$$
$$\beta_5 \times X_{\text{growth (from 36 weeks)}} \beta_6 \times X_{\text{medication for PH}} + \beta_5 \times X_{\text{non-invasive positive pressure}} +$$
$$\beta_6 \times X_{\text{invasive positive pressure}} + \beta_7 \times X_{\text{PEEP}}$$

for the 44 week time points.

Table 4 delineates the coefficients derived from the mixed effects models. The introduction of the random effect has induced alterations in the coefficients compared to the fixed effects models. Notably, for variables such as small for gestational age (SGA), medication for pulmonary hypertension (PH), and non-invasive positive pressure (exclusive to the 44-week model), the impact on the log odds of tracheostomy is now negative, contrasting with the positive effects observed in the fixed effects models.

Table 4: Coefficients for Random Effect Models: 36 Weeks (Left) and 44 Weeks (Right)

	Value		Value
Prenatal Corticosteroids	0.1271668	Cesarean Section	0.0118742
SGA	-0.0930739	Prenatal Corticosteroids	0.1819534
Growth (from Birth)	-0.0316410	Maternal Chorioamnionitis	0.0358817
Inspired O2	0.1026915	SGA	-0.0391717
Medication for PH	-0.0568918	Growth (from 36 Weeks)	0.0092892
Non-Invasive Positive Pressure	0.0427656	Medication for PH	-0.0300143
Invasive Positive Pressure	0.2328862	Non-Invasive Positive Pressure	-0.1223099
		Invasive Positive Pressure	0.2661314
		PEEP	0.0936958

Model Performance

Model evaluation encompassed precision, recall, and the F-score, with the optimal threshold determined by the ROC curve and documented in Table 5. Notably, a discernible trend suggests higher predictive accuracy in the 44-week model compared to the 36-week model. This observation extends to improved precision and recall metrics, coupled with a higher optimal threshold for predictions in the 44-week model.

Table 6 delineates the mixed effects model's performance on validation data. Strikingly, the introduction of the random intercept yields no significant change in model performance. This outcome suggests that

the anticipated substantial variation between centers might be less pronounced than initially envisaged. To further explore this, the Intra-Class Correlation Coefficient (ICC) was computed using the formula:

$$ICC = \frac{\sigma^2_{center}}{\sigma^2 + \sigma^2_{center}}$$

where σ^2_{center} captures between center variability and the denominator captures total variability.

Calculated ICCs for the 36-week and 44-week mixed effects models stand at 0.2022 and 0.2226, respectively. To justify the adoption of a random intercept, an ICC greater than 0.5 is typically sought, indicating substantial within-group variability amenable to capture by a random effect.

Table 5: Model Validation Metrics for Fixed Effect 36 and 44 Week Models

	Threshold	Precision	Recall	F Score
Week 36 Model	0.0928588	0.3826591	0.8838710	0.5340772
Week 44 Model	0.1875298	0.5078431	0.8516129	0.6361978

Table 6: Model Validation Metrics for Mixed Effect 36 and 44 Week Models

	Threshold	Precision	Recall	F Score
Week 36 Model	0.1832702	0.4268293	0.9032258	0.5797101
Week 44 Model	0.3826176	0.5757576	0.8580645	0.6891192

Discussion

Figure 2 illustrates the plotted predicted probabilities corresponding to variables common to both fixed effects models, namely ventilation support level, small for gestational age (SGA), medication for pulmonary hypertension (PH), and prenatal corticosteroids. Examination of the distributions reveals that, for most covariates, differences are not overtly pronounced. In the instances of prenatal corticosteroids and SGA, variations are discernible in the spread of distributions, rather than the location of their peaks. Conversely, medication for PH exhibits a more dispersed distribution among individuals receiving the medication, lacking distinct peaks.

Conversely, a notable distinction emerges in the predicted probabilities between infants devoid of supplementary oxygen and those subjected to invasive positive pressure. The probability distribution for infants without supplementary oxygen is confined, with a peak near 0, whereas the distribution for infants receiving invasive positive pressure is more expansive, featuring a peak closer to 0.4.

Limitations

Despite the primary objective of constructing a predictive model for tracheostomy decisions in infants with sBPD, several factors impede the model's efficacy. The notably small sample size, comprising less than 1,000 observations, poses a limitation, demanding an augmentation of training data to enhance variance capture. Furthermore, the data emanates from 10 distinct centers, yet crucial information about their characteristics, locations, referral status, or patient demographics remains undisclosed. The absence of such details hampers the assertion that the training data faithfully mirrors the demographic profile of the broader U.S. infant population with sBPD.

The misclassification of the race column raises concerns about the potential oversight of a race effect. While acknowledging the absence of a biological basis for race, potential associations with medical conditions like hypertension may exist, impacting pregnancy outcomes.

7

The issue of missing data, despite partial imputation, introduces uncertainty regarding imputation quality. Notably, the inability to impute the surfactant variable, a potentially pivotal predictor, which was excluded from variable selection.

The apparent class imbalance in the outcome variable further constrains the model, with a disproportionately higher number of cases where infants did not undergo tracheostomy. In an ideal scenario, balanced data would be preferred, and the Synthetic Minority Oversampling Technique (SMOTE) could be contemplated as a corrective measure to address this imbalance. Overall, these considerations underscore the need for meticulous scrutiny and refinement to bolster the model's reliability and generalizability for clinical decision-making.

Assumptions

The model operates under the assumption of observation independence, a concern mitigated by segregating measurements from the 36-week and 44-week intervals. Additionally, efforts have been directed at addressing potential severe multicollinearity by excluding variables pertaining to birth length and birth head circumference.

An acknowledgment of outliers in the dataset is imperative. However, given the inherently diverse and clinically complex nature of premature infants encompassing varying degrees of illness, the presence of variability in predictor variables is anticipated and contextually justified.

Conclusion

In conclusion, this study sheds light on the intricate landscape of predicting tracheostomy in infants with sBPD. The identification of key predictors, such as prenatal corticosteroids, SGA, inspired O2, medication for pulmonary hypertension (PH), and invasive positive pressure, underscores the multifaceted nature of the predictive model. In practice, the 44-week model demonstrates superior predictive accuracy, precision, and recall, emphasizing the importance of considering temporal dynamics in this analysis. This study contributes valuable insights into the challenges of predicting tracheostomy in sBPD infants. The current model exhibits potential for refinement, offering avenues for advancements in future research within this domain. The implementation of SMOTE could effectively address the existing class imbalance in the dataset. Furthermore, enhancing the model's predictive capacity could be facilitated by the acquisition of supplementary data pertaining to premature infants diagnosed with sBPD. These proposed enhancements hold promise for augmenting the robustness and efficacy of the predictive model in subsequent investigations.
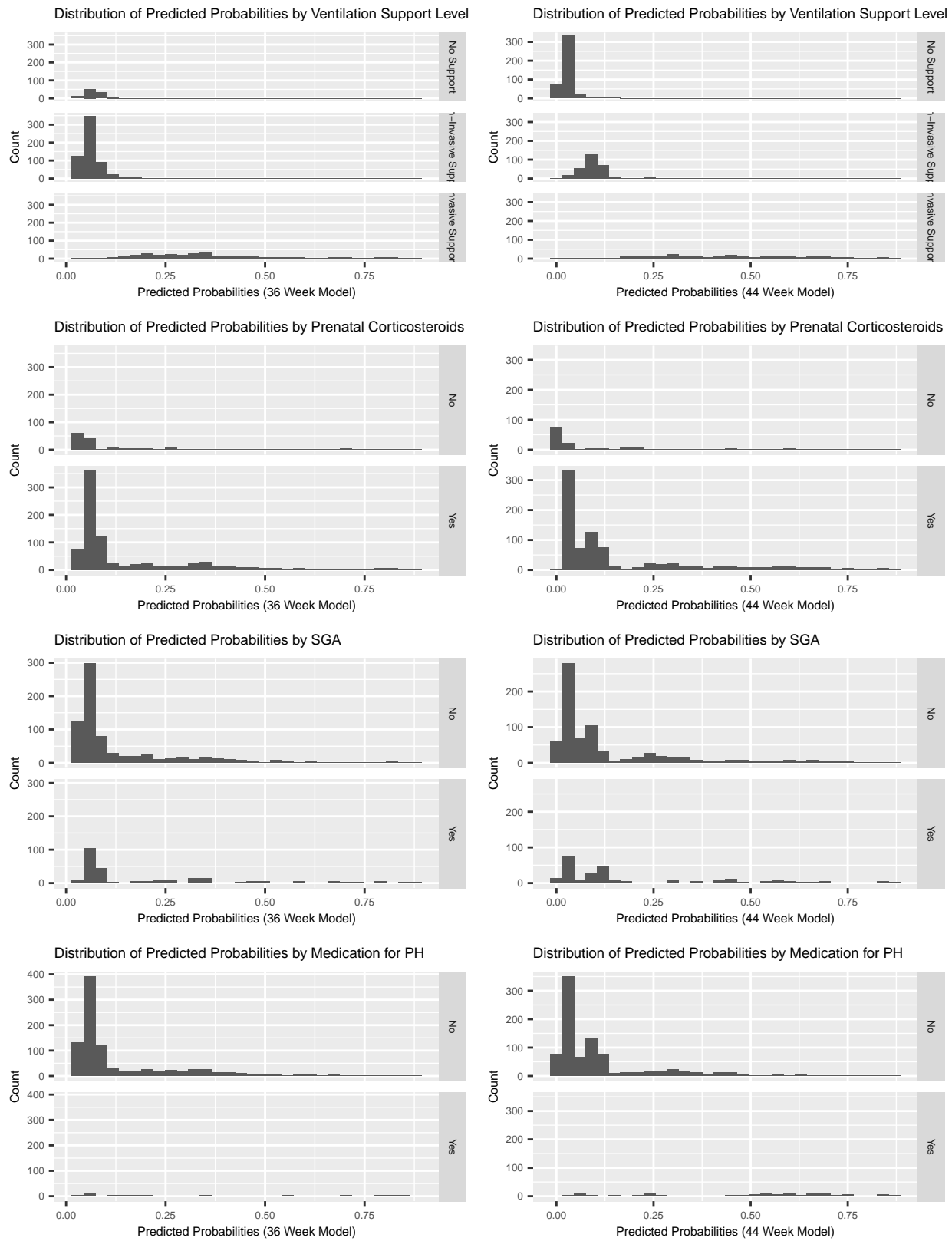
Figure 2. Plot of Predicted Probabilities by Ventilation Status

Code Appendix:

```r
knitr::opts_chunk$set(echo = FALSE, warning = FALSE,
    message = FALSE)
# Pre-processing Load in packages
library(tidyverse)
library(gtsummary)
library(kableExtra)
library(corrplot)
library(mice)
suppressPackageStartupMessages(library(glmnet))
suppressPackageStartupMessages(library(bestglm))
library(caret)
library(abess)
library(pROC)
library(gt)
library(lme4)


data <- read.csv("../project2.csv", stringsAsFactors = T)


# Data cleaning There is one record_id with 4
# identical rows. I remove the three duplicates
# here.
data <- data %>%
    distinct()


data[is.na(data$center), "center"] <- 1


# Remove record id and race Create variables for
# c section, hispanic/latino, male, and sga for
# the summary table to be more interpretable
data <- data %>%
    dplyr::select(-c(record_id, mat_race)) %>%
    mutate(csection = case_when(del_method == 2 ~ 1,
        del_method == 1 ~ 0, TRUE ~ NA), hisp_lat = case_when(mat_ethn ==
        1 ~ 1, mat_ethn == 2 ~ 0, TRUE ~ NA), male = case_when(gender ==
        "Female" ~ 0, gender == "Male" ~ 1, TRUE ~
        NA), sga = case_when(sga == "SGA" ~ 1, sga ==
        "Not SGA" ~ 0, TRUE ~ NA))
# Exploratory Data Analysis Summary Table by
# Tracheostomy Includes p-values to comment on
# differences between the groups
gt_table <- data %>%
    mutate(ventilation_support_level.36 = case_when(ventilation_support_level.36 ==
        0 ~ "No Respiratory Support", ventilation_support_level.36 ==
        1 ~ "Non-Invasive Positive Pressure", ventilation_support_level.36 ==
        2 ~ "Invasive Positive Pressure", TRUE ~ NA),
        ventilation_support_level_modified.44 = case_when(ventilation_support_level_modified.44 ==
            0 ~ "No Respiratory Support", ventilation_support_level_modified.44 ==
            1 ~ "Non-Invasive Positive Pressure", ventilation_support_level_modified.44 ==
            2 ~ "Invasive Positive Pressure", TRUE ~
            NA), Trach = case_when(Trach == 1 ~ "Yes",
            Trach == 0 ~ "No")) %>%
    tbl_summary(by = "Trach", missing = "no", include = c("hisp_lat",
```

```r
            "bw", "csection", "prenat_ster", "com_prenat_ster",
            "mat_chorio", "male", "sga", "any_surf", "weight_today.36",
            "ventilation_support_level.36", "inspired_oxygen.36",
            "p_delta.36", "peep_cm_h2o_modified.36", "med_ph.36",
            "weight_today.44", "ventilation_support_level_modified.44",
            "inspired_oxygen.44", "p_delta.44", "peep_cm_h2o_modified.44",
            "med_ph.44", "hosp_dc_ga"), label = list(hisp_lat ~
            "Hispanic or Latino Mother", bw ~ "Birth Weight",
            csection ~ "Cesarean Delivery", prenat_ster ~
                "Prenatal Corticosteroids", com_prenat_ster ~
                "Complete Prenatal Steroids", mat_chorio ~
                "Maternal Chorioamnionitis", male ~ "Male",
            sga ~ "Small for Gestational Age", any_surf ~
                "Any Surfactant", weight_today.36 ~ "Weight (36 Weeks)",
            ventilation_support_level.36 ~ "Ventilation Support Level (36 Weeks)",
            inspired_oxygen.36 ~ "Inspired O2 (36 Weeks)",
            p_delta.36 ~ "Peak Inspiratory Pressure (36 Weeks)",
            peep_cm_h2o_modified.36 ~ "Positive and Exploratory Pressure (36 Weeks)",
            med_ph.36 ~ "Medication for Pulmonary Hypertension (36 Weeks)",
            weight_today.44 ~ "Weight (44 Weeks)", ventilation_support_level_modified.44 ~
                "Ventilation Support Level (44 Weeks)",
            inspired_oxygen.44 ~ "Inspired O2 (44 Weeks)",
            p_delta.44 ~ "Peak Inspiratory Pressure (44 Weeks)",
            peep_cm_h2o_modified.44 ~ "Positive and Exploratory Pressure (44 Weeks)",
            med_ph.44 ~ "Medication for Pulmonary Hypertension (44 Weeks)",
            hosp_dc_ga ~ "Hospital Discharge Gestational Age")) %>%
    add_p() %>%
    as_gt()

tab_header(gt_table, title = "Table 1. Differences by Tracheostomy Status")
# Multicollinearity Create a correlation plot of
# numeric variables
cor_mat <- data %>%
    select_if(is.numeric) %>%
    dplyr::select(bw, ga, blength, birth_hc, weight_today.36,
        inspired_oxygen.36, p_delta.36, peep_cm_h2o_modified.36,
        weight_today.44, inspired_oxygen.44, p_delta.44,
        peep_cm_h2o_modified.44) %>%
    cor(use = "complete.obs")

colnames(cor_mat) <- c("Birth Weight", "Gestational Age",
    "Birth Length", "Birth Head Circumference", "Weight (36)",
    "Inspired O2 (36)", "Peak Inspiratory Press
ure (36)",
    "PEEP (36)", "Weight (44)", "Inspired O2 (44)",
    "Peak Inspiratory Pressure (44)", "PEEP (44)")
rownames(cor_mat) <- c("Birth Weight", "Gestational Age",
    "Birth Length", "Birth Head Circumference", "Weight (36)",
    "Inspired O2 (36)", "Peak Inspiratory Pressure (36)",
    "PEEP (36)", "Weight (44)", "Inspired O2 (44)",
    "Peak Inspiratory Pressure (44)", "PEEP (44)")

corrplot(cor_mat, method = "circle", type = "lower",
```

```r
        tl.cex = 0.8, tl.srt = 45, tl.col = "black")


# corrplot(cor_mat, method = 'shade', type =
# 'lower', tl.cex = 0.8, tl.col = 'black', tl.srt
# = 45) Multiple Imputation Convert variables to
# factors
data$mat_ethn <- as.factor(data$mat_ethn)
data$del_method <- as.factor(data$del_method)
data$prenat_ster <- as.factor(data$prenat_ster)
data$com_prenat_ster <- as.factor(data$com_prenat_ster)
data$mat_chorio <- as.factor(data$mat_chorio)
data$gender <- as.factor(data$gender)
data$sga <- as.factor(data$sga)
data$any_surf <- as.factor(data$any_surf)
data$center <- as.factor(data$center)
data$ventilation_support_level.36 <- as.factor(data$ventilation_support_level.36)
data$ventilation_support_level_modified.44 <- as.factor(data$ventilation_support_level_modified.44)

# Create a table of the variables with the
# highest rates of missingness
prop_missing <- c()
for (i in colnames(data)) {
    prop_missing <- c(prop_missing, sum(is.na(data[,
        i]))/(length(data[, i])) * 100)
}

missing_data <- data.frame(column = colnames(data),
    percent_missing = round(prop_missing, 2))

missing_data <- missing_data %>%
    filter(percent_missing > 15) %>%
    arrange(desc(percent_missing))
missing_data$column <- c("Inspired Oxygen (44)", "Peak Inspiratory Pressure (44)",
    "Weight (44)", "PEEP (44)", "Any Surfactant", "Ventilation Support (44)",
    "Medication for PH (44)", "Complete Prenatal Steroids")
colnames(missing_data) <- c("Column", "Percent Missing")

kbl(missing_data, booktabs = T, escape = F, caption = "Columns with Largest Volume of Missing Data") %>%
    kable_styling(latex_options = "HOLD_position")
# Multiple Imputation Don't want to use center
# Don't want to use any_surf - such a high volume
# of missing data - 43% Can impute other
# variables

set.seed(1)
index <- createDataPartition(data$Trach, p = 0.2, list = FALSE)
ignore_rows <- rownames(data) %in% index

data_mice_out <- mice(data[, -c(1, 13, 28:31)], 5,
    pri = F, ignore = ignore_rows)

# Store each imputed data set
```

```r
train_imp <- vector("list", 5)
test_imp <- vector("list", 5)
for (i in 1:5) {
    train_imp[[i]] <- mice::complete(data_mice_out,
        i)[-index, ]
    train_imp[[i]]$center <- data$center[-index]

    train_imp[[i]] <- train_imp[[i]] %>%
        mutate(growth_36 = weight_today.36 - bw, growth_44 = weight_today.44 -
            weight_today.36)

    test_imp[[i]] <- mice::complete(data_mice_out,
        i)[index, ]
    test_imp[[i]]$center <- data$center[index]
    test_imp[[i]] <- test_imp[[i]] %>%
        mutate(growth_36 = weight_today.36 - bw, growth_44 = weight_today.44 -
            weight_today.36)

    train_imp[[i]] <- train_imp[[i]] %>%
        mutate_at(c("bw", "weight_today.36", "inspired_oxygen.36",
            "p_delta.36", "peep_cm_h2o_modified.36",
            "weight_today.44", "inspired_oxygen.44",
            "p_delta.44", "peep_cm_h2o_modified.44",
            "growth_36", "growth_44"), function(x) (x -
            mean(x))/sd(x))
    test_imp[[i]] <- test_imp[[i]] %>%
        mutate_at(c("bw", "weight_today.36", "inspired_oxygen.36",
            "p_delta.36", "peep_cm_h2o_modified.36",
            "weight_today.44", "inspired_oxygen.44",
            "p_delta.44", "peep_cm_h2o_modified.44",
            "growth_36", "growth_44"), function(x) (x -
            mean(x))/sd(x))
}
forward_selection <- function(df, colnames) {
    #' Function to perform forward selection
    #'
    #' @param df, dataframe
    #' @param colnames, list of columns to consider in forward selection
    #'
    #' @return coefficients after performing k-fold cross validation

    x.ord <- model.matrix(Trach ~ ., data = df[, colnames])[,
        -1]
    y.ord <- df$Trach

    abess_fit <- abess(x.ord, y.ord, family = "binomial",
        tune.type = "cv", nfolds = 10, seed = 1)

    coefs <- extract(abess_fit, which.min(abess_fit$tune.value))

    return(coefs$beta)
}
```

```r
colnames_36 <- c("mat_ethn", "bw", "ga", "del_method",
    "prenat_ster", "com_prenat_ster", "mat_chorio",
    "gender", "sga", "growth_36", "inspired_oxygen.36",
    "p_delta.36", "peep_cm_h2o_modified.36", "med_ph.36",
    "ventilation_support_level.36", "Trach")


colnames_44 <- c("mat_ethn", "bw", "ga", "del_method",
    "prenat_ster", "com_prenat_ster", "mat_chorio",
    "gender", "sga", "growth_44", "inspired_oxygen.44",
    "p_delta.44", "peep_cm_h2o_modified.44", "med_ph.44",
    "ventilation_support_level_modified.44", "Trach")

# Perform forward selection for both the 36 week
# model and the 44 week model
coefs_1.36 <- forward_selection(train_imp[[1]], colnames_36)
coefs_2.36 <- forward_selection(train_imp[[2]], colnames_36)
coefs_3.36 <- forward_selection(train_imp[[3]], colnames_36)
coefs_4.36 <- forward_selection(train_imp[[4]], colnames_36)
coefs_5.36 <- forward_selection(train_imp[[5]], colnames_36)

coef_36 <- cbind(coefs_1.36, coefs_2.36, coefs_3.36,
    coefs_4.36, coefs_5.36)
avg_coefs_36 <- apply(coef_36, 1, mean)


coefs_1.44 <- forward_selection(train_imp[[1]], colnames_44)
coefs_2.44 <- forward_selection(train_imp[[2]], colnames_44)
coefs_3.44 <- forward_selection(train_imp[[3]], colnames_44)
coefs_4.44 <- forward_selection(train_imp[[4]], colnames_44)
coefs_5.44 <- forward_selection(train_imp[[5]], colnames_44)

coef_44 <- cbind(coefs_1.44, coefs_2.44, coefs_3.44,
    coefs_4.44, coefs_5.44)
avg_coefs_44 <- apply(coef_44, 1, mean)
# Add the random Effect Combine all test datasets
# into one big dataset for fitting the mixed
# effects model
test_data_long <- rbind(test_imp[[1]], test_imp[[2]],
    test_imp[[3]], test_imp[[4]], test_imp[[5]])

mixed.36 <- lmer(Trach ~ prenat_ster + sga + growth_36 +
    inspired_oxygen.36 + med_ph.36 + ventilation_support_level.36 +
    (1 | center), data = test_data_long)
predict.36 <- predict(mixed.36)

mixed.44 <- lmer(Trach ~ del_method + prenat_ster +
    mat_chorio + sga + growth_44 + med_ph.44 + ventilation_support_level_modified.44 +
    peep_cm_h2o_modified.44 + (1 | center), data = test_data_long)
predict.44 <- predict(mixed.44)
# Create tables containing coefficients for both
# fixed effects models
coefs_36_df <- as.data.frame(avg_coefs_36[avg_coefs_36 !=
```

```r
    0])
rownames(coefs_36_df) <- c("Prenatal Corticosteroids",
    "SGA", "Growth (from Birth)", "Inspired O2", "Medication for PH",
    "Invasive Positive Pressure")
colnames(coefs_36_df) <- c("Value")

coefs_44_df <- as.data.frame(avg_coefs_44[avg_coefs_44 !=
    0])
rownames(coefs_44_df) <- c("Cesarean Section", "Prenatal Corticosteroids",
    "Maternal Chorioamnionitis", "SGA", "Growth (from 36 Weeks)",
    "PEEP", "Medication for PH", "Invasive Positive Pressure")
colnames(coefs_44_df) <- c("Value")

kable(list(coefs_36_df, coefs_44_df), caption = "Coefficients for Fixed Effect Models: 36 Weeks (Left) a
    kable_styling(latex_options = "HOLD_position")
# Create tables containing coefficients for both
# mixed effects models
coef.mixed_36 <- as.data.frame(t(coef(mixed.36)$center[1,
    -1]))
colnames(coef.mixed_36) <- c("Value")
rownames(coef.mixed_36) <- c("Prenatal Corticosteroids",
    "SGA", "Growth (from Birth)", "Inspired O2", "Medication for PH",
    "Non-Invasive Positive Pressure", "Invasive Positive Pressure")

coef.mixed_44 <- as.data.frame(t(coef(mixed.44)$center[1,
    -1]))
colnames(coef.mixed_44) <- c("Value")
rownames(coef.mixed_44) <- c("Cesarean Section", "Prenatal Corticosteroids",
    "Maternal Chorioamnionitis", "SGA", "Growth (from 36 Weeks)",
    "Medication for PH", "Non-Invasive Positive Pressure",
    "Invasive Positive Pressure", "PEEP")

kable(list(coef.mixed_36, coef.mixed_44), caption = "Coefficients for Random Effect Models: 36 Weeks (Le
    kable_styling(latex_options = "HOLD_position")
# Model Validation Use ROC curve to get the best
# threshold

get_metrics <- function(data, model_coefs) {
    #' Function to get metrics for a model
    #'
    #' @param data, dataset
    #' @param model_coefs, named list of model coefficients
    #'
    #' @return list containing best threshold, precision, recall, and f score

    x_vars <- model.matrix(Trach ~ ., data = data)
    x_vars <- x_vars[, names(model_coefs)]

    data$score <- as.numeric(x_vars %*% model_coefs)
    model <- glm(Trach ~ score, data = data, family = "binomial")
    predictions <- predict(model, type = "response")
    best_threshold <- coords(roc(data$Trach ~ predictions,
        plot = FALSE, print.auc = FALSE), "best", ret = "threshold")[[1]]
```

```r
    predictions <- as.numeric(predict(model, type = "response") >
        best_threshold)

    true <- data$Trach
    pred <- predictions

    # Calculate Precision and Recall
    tp = fp = tn = fn = 0
    for (i in 1:length(predictions)) {
        if (pred[i] == 1 & true[i] == 1) {
            tp <- tp + 1
        } else if (pred[i] == 0 & true[i] == 0) {
            tn <- tn + 1
        } else if (pred[i] == 1 & true[i] == 0) {
            fp <- fp + 1
        } else {
            fn <- fn + 1
        }
    }

    precision <- tp/(tp + fp)
    recall <- tp/(tp + fn)

    f1 <- ((recall^(-1) + precision^(-1))/2)^(-1)

    return(list(best_threshold = best_threshold, precision = precision,
        recall = recall, f1 = f1))

}


model_validation <- function(test_data, coefs) {
    #' Function to loop through all imputed test datasets and calculate the metrics for each one
    #' Averages the metrics across all imputed datasets
    #'
    #' @param test_data, list of imputed test datasets
    #' @param coefs, named list of coefficients
    #'
    #' @return list containing average values for all metrics across all imputed datasets

    precision = recall = f1 = threshold = c()

    for (i in 1:length(test_data)) {

        metrics <- get_metrics(test_data[[i]], coefs)
        precision <- c(precision, metrics$precision)
        recall <- c(recall, metrics$recall)
        f1 <- c(f1, metrics$f1)
        threshold <- c(threshold, metrics$best_threshold)

    }

    return(list(threshold = mean(threshold), precision = mean(precision),
```

```r
        recall = mean(recall), f1 = mean(f1)))

}


# Perform model validation for fixed effects
# models
val.36 <- model_validation(test_imp, avg_coefs_36)
val.44 <- model_validation(test_imp, avg_coefs_44)


val_df <- t(data.frame(week36 = as.numeric(val.36),
    week44 = as.numeric(val.44)))
rownames(val_df) <- c("Week 36 Model", "Week 44 Model")
colnames(val_df) <- c("Threshold", "Precision", "Recall",
    "F Score")
kbl(val_df, caption = "Model Validation Metrics for Fixed Effect 36 and 44 Week Models",
    booktabs = T, escape = F) %>%
    kable_styling(latex_options = "HOLD_position")


model_val <- function(data, model) {
    #' Function to get metrics for a model
    #'
    #' @param data, dataset
    #' @param model, model object
    #'
    #' @return list containing best threshold, precision, recall, and f score

    predictions <- predict(model, type = "response")
    best_threshold <- coords(roc(data$Trach ~ predictions,
        plot = FALSE, print.auc = FALSE), "best", ret = "threshold")[[1]]

    true <- data$Trach
    predictions <- as.numeric(predict(model, type = "response") >
        best_threshold)

    # Calculate Precision and Recall
    tp = fp = tn = fn = 0
    for (i in 1:length(predictions)) {
        if (predictions[i] == 1 & true[i] == 1) {
            tp <- tp + 1
        } else if (predictions[i] == 0 & true[i] == 0) {
            tn <- tn + 1
        } else if (predictions[i] == 1 & true[i] == 0) {
            fp <- fp + 1
        } else {
            fn <- fn + 1
        }
    }

    precision <- tp/(tp + fp)
    recall <- tp/(tp + fn)
```

```r
    f1 <- ((recall^(-1) + precision^(-1))/2)^(-1)

    return(list(best_threshold = best_threshold, precision = precision,
        recall = recall, f1 = f1))

}

# Perform model validation for mixed effects
# models
mixed_val.36 <- model_val(test_data_long, mixed.36)
mixed_val.44 <- model_val(test_data_long, mixed.44)

mixed_val_df <- t(data.frame(week36 = as.numeric(mixed_val.36),
    week44 = as.numeric(mixed_val.44)))
rownames(mixed_val_df) <- c("Week 36 Model", "Week 44 Model")
colnames(mixed_val_df) <- c("Threshold", "Precision",
    "Recall", "F Score")
kbl(mixed_val_df, caption = "Model Validation Metrics for Mixed Effect 36 and 44 Week Models",
    booktabs = T, escape = F) %>%
    kable_styling(latex_options = "HOLD_position")
# Plot distributions of predicted probabilities
# by significant covariates

library(gridExtra)
x_vars <- model.matrix(Trach ~ ., data = test_data_long)
x_vars_44 <- x_vars[, names(avg_coefs_44)]
x_vars_36 <- x_vars[, names(avg_coefs_36)]

test_data_long$score_36 <- as.numeric(x_vars_36 %*%
    avg_coefs_36)
test_data_long$score_44 <- as.numeric(x_vars_44 %*%
    avg_coefs_44)

model_36 <- glm(Trach ~ score_36, data = test_data_long,
    family = "binomial")
model_44 <- glm(Trach ~ score_44, data = test_data_long,
    family = "binomial")

pred_36 <- predict(model_36, type = "response")
pred_44 <- predict(model_44, type = "response")

vent_levels <- c(`0` = "No Support", `1` = "Non-Invasive Support",
    `2` = "Invasive Support")

plot_1 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_36)) +
    facet_grid(rows = vars(ventilation_support_level.36),
        labeller = as_labeller(vent_levels)) + labs(title = "Distribution of Predicted Probabilities by
    x = "Predicted Probabilities (36 Week Model)",
    y = "Count") + theme(text = element_text(size = 6))

plot_2 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_44)) +
    facet_grid(rows = vars(ventilation_support_level_modified.44),
        labeller = as_labeller(vent_levels)) + labs(title = "Distribution of Predicted Probabilities by
```

```r
    x = "Predicted Probabilities (44 Week Model)",
    y = "Count") + theme(text = element_text(size = 6))


plot_3 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_36)) +
    facet_grid(rows = vars(prenat_ster)) + labs(title = "Distribution of Predicted Probabilities by Pren
    x = "Predicted Probabilities (36 Week Model)",
    y = "Count") + theme(text = element_text(size = 6))

plot_4 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_44)) +
    facet_grid(rows = vars(prenat_ster)) + labs(title = "Distribution of Predicted Probabilities by Pren
    x = "Predicted Probabilities (44 Week Model)",
    y = "Count") + theme(text = element_text(size = 6))

sga_levels <- c(`0` = "No", `1` = "Yes")

plot_5 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_36)) +
    facet_grid(rows = vars(sga), labeller = as_labeller(sga_levels)) +
    labs(title = "Distribution of Predicted Probabilities by SGA",
        x = "Predicted Probabilities (36 Week Model)",
        y = "Count") + theme(text = element_text(size = 6))

plot_6 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_44)) +
    facet_grid(rows = vars(sga), labeller = as_labeller(sga_levels)) +
    labs(title = "Distribution of Predicted Probabilities by SGA",
        x = "Predicted Probabilities (44 Week Model)",
        y = "Count") + theme(text = element_text(size = 6))

med_ph_levels <- c(`0` = "No", `1` = "Yes")

plot_7 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_36)) +
    facet_grid(rows = vars(med_ph.36), labeller = as_labeller(med_ph_levels)) +
    labs(title = "Distribution of Predicted Probabilities by Medication for PH",
        x = "Predicted Probabilities (36 Week Model)",
        y = "Count") + theme(text = element_text(size = 6))

plot_8 <- ggplot(test_data_long) + geom_histogram(aes(x = pred_44)) +
    facet_grid(rows = vars(med_ph.44), labeller = as_labeller(med_ph_levels)) +
    labs(title = "Distribution of Predicted Probabilities by Medication for PH",
        x = "Predicted Probabilities (44 Week Model)",
        y = "Count") + theme(text = element_text(size = 6))


grid.arrange(plot_1, plot_2, plot_3, plot_4, plot_5,
    plot_6, plot_7, plot_8, nrow = 4, ncol = 2)
```

References

n.d. *What Is Bronchopulmonary Dysplasia?* Nationwide Children's. https://www.nationwidechildrens.org/conditions/bronchopulmonary-dysplasia-bpd.

Akangire, Gangaram, and Winston Manimtim. 2023. "Tracheostomy in Infants with Severe Bronchopulmonary Dysplasia: A Review." *Frontiers in Pediatrics* 10: 1066367.

Hastie, Trevor, Robert Tibshirani, and R Tibshirani. 2020. "Best Subset, Forward Stepwise or Lasso." *Analysis and Recommendations Based on Extensive Comparisons: Statistical Science.*