A Prognostic Model for Tracheostomy and Mortality in Infants with Severe Bronchopulmonary Dysplasia at Different Stages

Monica Colon-Vargas

Fall 2023

Objective: The ideal criteria and timing for tracheostomy in severe bronchopulmonary dysplasia (sBPD) lack clarity, despite studies suggesting potential growth advantages with earlier placement. Past assessments using vast databases have shown precise forecasts of tracheostomy or mortality, relying on initial demographics and clinical diagnoses. This study aims to construct predictive models at three crucial stages in infants with sBPD: Birth, 36 weeks, and 44 weeks, while comparing the performance of a mixed model with StepAIC and a LASSO model.
Methods: This study employs two statistical techniques for each dataset, initially fitting a generalized mixed model to account for the nested data structure. Variable selection is carried out using the stepwise Akaike Information Criterion (stepAIC) algorithm. This process is repeated independently for each training imputed dataset, ensuring a comprehensive approach. Subsequently, a Lasso logistic known for reliable predictions and feature selection, is applied. Cross-validation determines the optimal penalization parameter, and the final model is derived by pooling coefficients from each training imputed dataset. This methodology is systematically applied across six distinct models for a comprehensive comparative analysis.
Conclusions: Mixed model outperformed LASSO model possibly due to LASSo not capturing the nested structure of the data concerning the centers. The use of prenatal steroids emerged as significant among all six models. As expected, the 36-week and 44-week models highlighted the importance of variables at that time, including weight, ventilation support, and medication for pulmonary hypertension. Model selection prioritized AUC scores, leading to recommending the use of Mixed models at each time point.

1 Introduction

The optimal criteria and timing for tracheostomy placement in neonates with severe bronchopulmonary dysplasia (sBPD) remain uncertain. Some studies propose potential benefits in growth with earlier tracheostomy placement. Previous examinations of extensive databases have demonstrated accurate prediction of tracheostomy placement or mortality based on baseline demographics and clinical diagnoses. However, these analyses have not incorporated detailed respiratory parameters and have not offered predictions at different postmenstrual ages (PMA). A precise prediction of the likelihood of tracheostomy at early PMA could significantly influence counseling for families and the timing of tracheostomy placement, a topic currently under active debate in the context of sBPD.

The primary objective of this study is to develop a regression model with the aim of predicting the composite outcome of tracheostomy or death. The utilization of a composite outcome in this study, encompassing both tracheostomy and death as endpoints, serves to offer a more comprehensive perspective on patient outcomes. This combined endpoint is of particular interest due to its clinical relevance and statistical efficiency. By incorporating multiple critical events into one outcome, the research can effectively capture the severity of the condition and the associated risks. The overarching goal is to contribute valuable insights that can guide the establishment of indication criteria and inform decisions regarding the optimal timing for

tracheostomy placement in neonates with severe sBPD. This model seeks to elevate the accuracy of prognostic assessments, ultimately enabling healthcare professionals to make more informed clinical decisions in managing the patient population with sBPD.

We chose not to include transformations or interactions to maintain clarity and ensure straightforward interpretation of our results. Simplifying the model helped us focus on the most essential aspects of the data, aligning with our goal of providing clear and reliable insights. Additionally, no theoretical or empirical basis suggesting specific interactions or transformations, including them might were not found.

2 Data and Methodology

The study participants were selected from the BPD Collaborative Registry, which is a collaborative initiative comprising interdisciplinary BPD programs in the United States and Sweden. The registry specifically includes infants born with a gestational age of less than 32 weeks and diagnosed with severe bronchopulmonary dysplasia (sBPD) based on the 2001 NHLBI criteria. In the registry, standard demographic and clinical data are collected at four-time points: birth, 36 weeks PMA, 44 weeks PMA and discharge.The data consists of individuals with BPD and comprehensive growth data recorded from January 1 to July 19, 2021 from 10 BPD Collaborative centers.

Data inaccuracies were identified, including instances of repeated identification, which were subsequently rectified by eliminating duplications. Additionally, due to the limited number of individuals in centers 20 and 21, they were consolidated into a unified category for analysis. Due to discrepancies in the levels of the maternal race, the variable `mat_race` was excluded from the analysis. The variable `any_surf` was removed due to the high amount of missing data.

Table 1 presents a comprehensive set of summary statistics for the dataset, with a focus on the two response variables, namely Tracheostomy and Death. It is essential to dissect these statistics within the context of various aspects related to the data. The analysis of mothers' ethnicity reveals a balanced distribution across both outcomes. A noticeable trend in the data is the lower birth weight of infants who either died or underwent a tracheostomy, which may indicate that low birth weight is a risk factor for these outcomes. Additionally, a majority of these births occurred through cesarean section. Furthermore, it is worth highlighting that males appear to be more affected, with a higher number of deaths and tracheostomies. This gender disparity warrants a more in-depth analysis to understand the underlying factors contributing to this observed difference in outcomes between males and females. Additionally, Table 2 presents the summary statistics for birth data and some 36-week data for all centers. The data appears to originate from ten different medical centers. It is noteworthy that center 20 (class containing both centers 20 and 21) contains relatively fewer cases, suggesting that they might be smaller facilities, in contrast to center 2, which evidently stands out as a larger center. Furthermore, center 2 is notably the primary provider of tracheostomies, closely followed by centers 12, 1, and 4. A similar trend is observed for the occurrence of deaths, with centers 2, 12, and 1 featuring more.

This study is structured into three distinct segments. Initially, we construct a predictive model for the composite outcome in newborn babies, utilizing a subset of the dataset that exclusively incorporates variables measured at birth (Birth-model). Subsequently, we develop a second model to forecast the necessity of a tracheotomy, taking into account variables assessed at the 36-week stage of the pregnancy (36-week model). In developing this model, we restrict our analysis to patients who possess at least one recorded measurement within the 36-week window. Lastly, the third model encompasses the complete dataset and is designed to

predict the occurrence of tracheotomies at the 44-week (44-week model). Consistent with the previous model, this dataset includes patients who have at least one recorded measurement within the 44-week window. This approach ensures that the models are developed and validated based on a comprehensive and appropriate population. The primary aim is to construct a predictive model designed to assess the probability of tracheostomy necessity at three distinct temporal moments in a patient's early life—specifically, at birth, 36 weeks, and 44 weeks. This predictive framework seeks to furnish anticipatory insights into the potential requirement for tracheostomy intervention across these critical time points.

For each dataset utilized in constructing an individual model, two statistical techniques are employed. Given the nested structure of the data, where each patient is nested within a center, a generalized mixed model is initially fitted, incorporating random intercepts with the center acting as a random effect. Considering the correlations among variables, such as the high correlation between birth weight (`bw`), birth length (`blength`), and birth head circumference (`birth_hc`), variable selection becomes imperative. The stepwise Akaike Information Criterion (stepAIC) algorithm is applied for this purpose, aiming to minimize the stepAIC value and determine the optimal set of features. We prioritize choosing variables carefully rather than using complex regression methods. We aim for models that are easier to understand, focusing on clear interpretations. By selecting variables smartly, we capture the most important predictors, giving us a concise but insightful view of the data's relationships. This algorithm is executed independently for each training imputed dataset (totaling 5), and the frequency of variable selection is recorded. Selecting all variables that have been chosen at least once for fitting compromised the model's ability to accurately capture the genuine significance of the variables. To avoid this, we adopt a criterion where only variables selected in more than 40% of instances are retained. Subsequently, the model is fitted to each training imputed dataset using the selected variables, and the resulting coefficients are aggregated through averaging. This rigorous process ensures that the final model captures the genuine significance of the selected variables while accommodating the nuances of the data structure. Subsequently, a Lasso logistic model is applied to the dataset, maintaining the center as a fixed effect. Despite the inherent limitation of not explicitly accommodating the nested structure of the data, this approach is pursued to facilitate a comparative analysis. The choice to employ Lasso in this context is grounded in its established reputation for delivering reliable predictions and its feature selection. To determine the optimal penalization parameter, cross-validation is performed independently on each training imputed dataset. Following the selection of coefficients for each training imputed dataset, pooling together these coefficients we derive the final model. In this stage, we factor in all the chosen coefficients, differing from the prior approach where only coefficients selected in more than 40% were considered. It is noteworthy to highlight that the described methodology is systematically reiterated for each model, culminating in a comparative assessment encompassing a total of six distinct models.

For model evaluation, we present results encompassing metrics such as the Area Under the Receiver Operating Characteristic Curve (AUC), Brier score, sensitivity, and specificity. Imputation procedures yielded five imputed datasets, each divided into training and testing sets. Models were trained using the training data and subsequently, their performance was evaluated on the testing sets. The metrics results were pooled together by averaging them. These metrics collectively provide an evaluation of the models' discriminative ability, calibration, and performance in correctly identifying true positive and true negative instances.

Table 1: Summary Statistics Grouped by Response Outcome

2*Variable	Death				Tracheostomy			
	N	Overall	No	Yes	N	Overall	No	Yes
mat_ethn	936				938			
1		74.00 (7.91%)	71.00 (8.02%)	3.00 (5.88%)		74.00 (7.89%)	66.00 (8.23%)	8.00 (5.88%)
2		862.00 (92.09%)	814.00 (91.98%)	48.00 (94.12%)		864.00 (92.11%)	736.00 (91.77%)	128.00 (94.12%)
bw	993	806.20 (296.98)	811.12 (288.99)	720.69 (406.03)	995	806.10 (296.77)	813.82 (295.25)	761.21 (302.58)
ga	993	25.77 (2.14)	25.77 (2.14)	25.83 (2.13)	995	25.77 (2.14)	25.76 (2.14)	25.84 (2.13)
blength	915	32.49 (3.82)	32.60 (3.75)	30.32 (4.37)	917	32.49 (3.82)	32.56 (3.80)	31.97 (3.94)
birth_hc	916	23.19 (2.76)	23.24 (2.71)	22.34 (3.54)	918	23.19 (2.76)	23.22 (2.71)	22.99 (3.07)
del_method	990				992			
1		284.00 (28.69%)	274.00 (29.27%)	10.00 (18.52%)		285.00 (28.73%)	254.00 (29.99%)	31.00 (21.38%)
2		706.00 (71.31%)	662.00 (70.73%)	44.00 (81.48%)		707.00 (71.27%)	593.00 (70.01%)	114.00 (78.62%)
prenat_ster	958	832.00 (86.85%)	788.00 (86.69%)	44.00 (89.80%)	960	834.00 (86.88%)	711.00 (85.77%)	123.00 (93.89%)
com_prenat_ster	922	607.00 (65.84%)	573.00 (65.56%)	34.00 (70.83%)	924	609.00 (65.91%)	523.00 (65.05%)	86.00 (71.67%)
mat_chorio	931	160.00 (17.19%)	151.00 (17.25%)	9.00 (16.36%)	933	160.00 (17.17%)	144.00 (17.36%)	16.00 (14.55%)
gender	989			991				
Female		407.00 (41.15%)	390.00 (41.71%)	17.00 (31.48%)		408.00 (41.17%)	348.00 (41.18%)	60.00 (41.10%)
Male		582.00 (58.85%)	545.00 (58.29%)	37.00 (68.52%)		583.00 (58.83%)	497.00 (58.82%)	86.00 (58.90%)
sga	978				980			
Not SGA		775.00 (79.24%)	747.00 (80.84%)	28.00 (51.85%)		777.00 (79.29%)	677.00 (80.79%)	100.00 (70.42%)
SGA		203.00 (20.76%)	177.00 (19.16%)	26.00 (48.15%)		203.00 (20.71%)	161.00 (19.21%)	42.00 (29.58%)
any_surf	561	460.00 (82.00%)	432.00 (81.66%)	28.00 (87.50%)	562	460.00 (81.85%)	394.00 (80.90%)	66.00 (88.00%)
weight_today.36	901	2,120.08 (413.67)	2,132.66 (397.30)	1,826.46 (633.79)	903	2,120.90 (413.58)	2,131.91 (410.26)	2,023.88 (432.03)
ventilation_support_level.36	963				965			
0		116.00 (12.05%)	114.00 (12.46%)	2.00 (4.17%)		117.00 (12.12%)	111.00 (13.25%)	6.00 (4.72%)
1		588.00 (61.06%)	581.00 (63.50%)	7.00 (14.58%)		588.00 (60.93%)	559.00 (66.71%)	29.00 (22.83%)
2		259.00 (26.90%)	220.00 (24.04%)	39.00 (81.25%)		260.00 (26.94%)	168.00 (20.05%)	92.00 (72.44%)
inspired_oxygen.36	901	0.34 (0.15)	0.33 (0.13)	0.52 (0.25)	903	0.34 (0.15)	0.32 (0.13)	0.49 (0.20)
p_delta.36	865	5.28 (9.75)	4.85 (9.37)	17.14 (12.70)	867	5.27 (9.74)	4.26 (8.87)	15.13 (12.13)
peep_cm_h2o_modified.36	876	6.34 (2.91)	6.31 (2.93)	7.13 (2.25)	878	6.33 (2.91)	6.20 (2.90)	7.50 (2.79)
med_ph.36	963				965			
0		898.00 (93.25%)	860.00 (93.99%)	38.00 (79.17%)		899.00 (93.16%)	797.00 (95.11%)	102.00 (80.31%)
1		65.00 (6.75%)	55.00 (6.01%)	10.00 (20.83%)		66.00 (6.84%)	41.00 (4.89%)	25.00 (19.69%)
weight_today.44	550	3,646.12 (682.09)	3,674.90 (662.68)	3,222.71 (822.51)	550	3,646.12 (682.09)	3,667.21 (665.21)	3,550.07 (750.40)
ventilation_support_level_modified.44	572				572			
0		269.00 (47.03%)	268.00 (50.28%)	1.00 (2.56%)		269.00 (47.03%)	262.00 (56.83%)	7.00 (6.31%)
1		146.00 (25.52%)	141.00 (26.45%)	5.00 (12.82%)		146.00 (25.52%)	128.00 (27.77%)	18.00 (16.22%)
2		157.00 (27.45%)	124.00 (23.26%)	33.00 (84.62%)		157.00 (27.45%)	71.00 (15.40%)	86.00 (77.48%)
inspired_oxygen.44	548	0.34 (0.15)	0.33 (0.13)	0.53 (0.23)	548	0.34 (0.15)	0.32 (0.13)	0.44 (0.19)
p_delta.44	548	7.62 (14.19)	6.44 (13.12)	26.75 (17.10)	548	7.62 (14.19)	4.80 (12.01)	21.10 (15.99)
peep_cm_h2o_modified.44	550	4.30 (4.46)	4.03 (4.40)	8.78 (2.71)	550	4.30 (4.46)	3.37 (4.06)	8.69 (3.59)
med_ph.44	572				572			
0		473.00 (82.69%)	458.00 (85.93%)	15.00 (38.46%)		473.00 (82.69%)	413.00 (89.59%)	60.00 (54.05%)
1		99.00 (17.31%)	75.00 (14.07%)	24.00 (61.54%)		99.00 (17.31%)	48.00 (10.41%)	51.00 (45.95%)
hosp_dc_ga	870	52.79 (26.54)	52.44 (26.66)	58.90 (23.75)	871	52.78 (26.53)	48.93 (23.63)	79.94 (29.94)
Trach	993	146.00 (14.70%)	129.00 (13.74%)	17.00 (31.48%)				
Death					993	54.00 (5.44%)	37.00 (4.37%)	17.00 (11.64%)

Table 2: Summary Statistics Grouped by Center

Characteristic	N	Overall, N = 985	1, N = 55	2, N = 629	3, N = 57	4, N = 60	5, N = 40	7, N = 32	12, N = 69	16, N = 38	20, N = 5
Death	983	54.00 (5.49%)	7.00 (12.73%)	29.00 (4.62%)	1.00 (1.75%)	1.00 (1.69%)	2.00 (5.00%)	0.00 (0.00%)	14.00 (20.29%)	0.00 (0.00%)	0.00 (0.00%)
Trach	985	142.00 (14.42%)	23.00 (41.82%)	64.00 (10.17%)	1.00 (1.75%)	11.00 (18.33%)	5.00 (12.50%)	1.00 (3.13%)	35.00 (50.72%)	1.00 (2.63%)	1.00 (20.00%)
bw	985	806.99 (296.65)	684.62 (196.95)	832.43 (312.86)	764.81 (254.94)	833.25 (258.78)	605.35 (107.03)	724.88 (225.73)	781.00 (253.55)	889.29 (354.17)	989.00 (422.35)
weight_today.36	898	2,121.62 (412.56)	2,072.56 (440.04)	2,135.40 (408.47)	2,105.93 (424.61)	2,126.17 (339.43)	1,921.83 (401.40)	2,169.32 (408.61)	2,039.93 (478.50)	2,219.50 (408.51)	2,241.80 (526.55)
del_method	983	-	284.00 (28.89%)	177.00 (28.14%)	17.00 (29.82%)	18.00 (30.00%)	14.00 (35.00%)	10.00 (31.25%)	18.00 (26.87%)	14.00 (36.84%)	1.00 (20.00%)
1	-	-	284.00 (28.89%)	177.00 (28.14%)	17.00 (29.82%)	18.00 (30.00%)	14.00 (35.00%)	10.00 (31.25%)	18.00 (26.87%)	14.00 (36.84%)	1.00 (20.00%)
2	-	-	699.00 (71.11%)	452.00 (71.86%)	40.00 (70.18%)	42.00 (70.00%)	26.00 (65.00%)	22.00 (68.75%)	49.00 (73.13%)	24.00 (63.16%)	4.00 (80.00%)
prenat_ster	950	824.00 (86.74%)	46.00 (90.20%)	543.00 (86.46%)	47.00 (87.04%)	47.00 (79.66%)	37.00 (92.50%)	26.00 (86.67%)	41.00 (89.13%)	33.00 (89.19%)	4.00 (80.00%)
ventilation_support_level.36	956	-	-	-	-	-	-	-	-	-	-
0	-	116.00 (12.13%)	7.00 (12.73%)	50.00 (8.06%)	5.00 (8.93%)	8.00 (13.33%)	0.00 (0.00%)	22.00 (68.75%)	1.00 (2.00%)	22.00 (57.89%)	1.00 (20.00%)
1	-	585.00 (61.19%)	19.00 (34.55%)	424.00 (68.39%)	35.00 (62.50%)	34.00 (56.67%)	31.00 (77.50%)	8.00 (25.00%)	17.00 (34.00%)	14.00 (36.84%)	3.00 (60.00%)
2	-	255.00 (26.67%)	29.00 (52.73%)	146.00 (23.55%)	16.00 (28.57%)	18.00 (30.00%)	9.00 (22.50%)	2.00 (6.25%)	32.00 (64.00%)	2.00 (5.26%)	1.00 (20.00%)
inspired_oxygen.36	897	0.34 (0.15)	0.43 (0.21)	0.32 (0.14)	0.32 (0.10)	0.40 (0.12)	0.36 (0.13)	0.36 (0.10)	0.40 (0.19)	0.35 (0.11)	0.41 (0.28)
p_delta.36	860	5.27 (9.76)	7.46 (8.40)	5.31 (10.78)	6.74 (7.72)	5.21 (5.73)	4.06 (6.74)	0.14 (0.79)	9.11 (7.17)	1.29 (4.67)	10.25 (7.93)
med_ph.36	956	-	-	-	-	-	-	-	-	-	-
0	-	890.00 (93.10%)	42.00 (76.36%)	595.00 (95.97%)	53.00 (94.64%)	49.00 (81.67%)	37.00 (92.50%)	30.00 (93.75%)	46.00 (92.00%)	34.00 (89.47%)	4.00 (80.00%)
1	-	66.00 (6.90%)	13.00 (23.64%)	25.00 (4.03%)	3.00 (5.36%)	11.00 (18.33%)	3.00 (7.50%)	2.00 (6.25%)	4.00 (8.00%)	4.00 (10.53%)	1.00 (20.00%)

2.1 Outcome Variable

The objective of this study is to create a predictive model for the combined outcomes of tracheotomy and death. This model will aid in determining when and for whom tracheotomy placement is appropriate, thus improving clinical decision-making in patient care. The initial step in this study involves the construction of a composite outcome variable. To delineate this composite outcome, the following approach is adopted: Patients who have undergone a tracheotomy, without consideration of their subsequent survival status,

are categorized into the "tracheotomy" group. Likewise, patients who did not receive a tracheotomy but subsequently deceased are also classified under the "tracheotomy" category. This particular categorization serves the purpose of recognizing cases where death might have been averted if a tracheotomy had been performed. Conversely, individuals who neither underwent a tracheotomy nor succumbed to mortality are categorized under the "no tracheotomy" group. This approach allows us to create a comprehensive analysis of the combined outcome involving tracheotomies and deaths. By considering cases where tracheotomies might have prevented deaths, we can gain a fuller understanding of the clinical decisions and preventive actions needed in this scenario. Following the specification of our outcome variable, the dataset exhibits a class imbalance with 810 instances in the negative class and 183 instances in the positive class. Given this imbalanced distribution, it is anticipated that the optimal threshold in probability for prediction will be constrained to lower values. This adjustment aims to achieve a balanced trade-off between sensitivity and specificity in the predictive model.

2.2 Missing Data

The dataset exhibits a notable issue with missing data, where approximately 15% of cases lack complete information. From Table 1, this discrepancy is primarily observed in variables collected at the 44-week mark, with a similar pattern evident in the variable denoting whether the baby received surfactant treatment within the first 72 hours of life `any_surf`. Notably, variables measured at 36 weeks also display substantial rates of missing data. Figure 1 illustrates the pattern of missing values in relation to the hospital discharge timing of patients. Notably, patients discharged prior to the 44-week mark exhibit a higher prevalence of missing values in variables collected at that specific time point. In contrast, patients discharged after the 44th week demonstrate a comparatively lower incidence of missing values in the same variables.

The three datasets utilized for constructing the three models exhibit a significant prevalence of missing values. From Figure 1, apart from data collected at 44 weeks, missing data pattern might imply data is missing at random. To enhance the model's quality and robustness, the adoption of a technique to address this issue becomes imperative. Multiple imputation is deployed across the three datasets, resulting in the creation of five complete datasets for each, aggregating to a total of 15 complete datasets. Subsequently, models are fitted to these complete datasets, yielding model estimates which are subsequently pooled, culminating in the derivation of three final model (one for each dataset).

Variable	n	%	Variable	n	%
inspired_oxygen	447	44.92	birth_hc	77	7.74
p_delta.44	447	44.92	com_prenat_ster	71	7.14
weight_today.44	445	44.72	mat_chorio	62	6.23
peep_cm_h2o_modified.44	445	44.72	mat_ethn	57	5.73
any_surf	433	43.52	prenat_ster	35	3.52
ventilation_support_level_modified.44	423	42.51	ventilation_support_level.36	30	3.02
med_ph.44	423	42.52	med_ph.36	30	3.02
p_delta.36	128	12.86	sga	15	1.51
hosp_dc_ga	124	12.46	center	10	1.01
peep_cm_h20_modified.36	117	11.76	gender	4	0.40
weight_today.36	92	9.25	del_method	3	0.30
inspired_oxygen.36	92	9.25	Death	2	0.20
blength	78	7.84	new_trach	2	0.20

Table 3: Missing Data per Variables

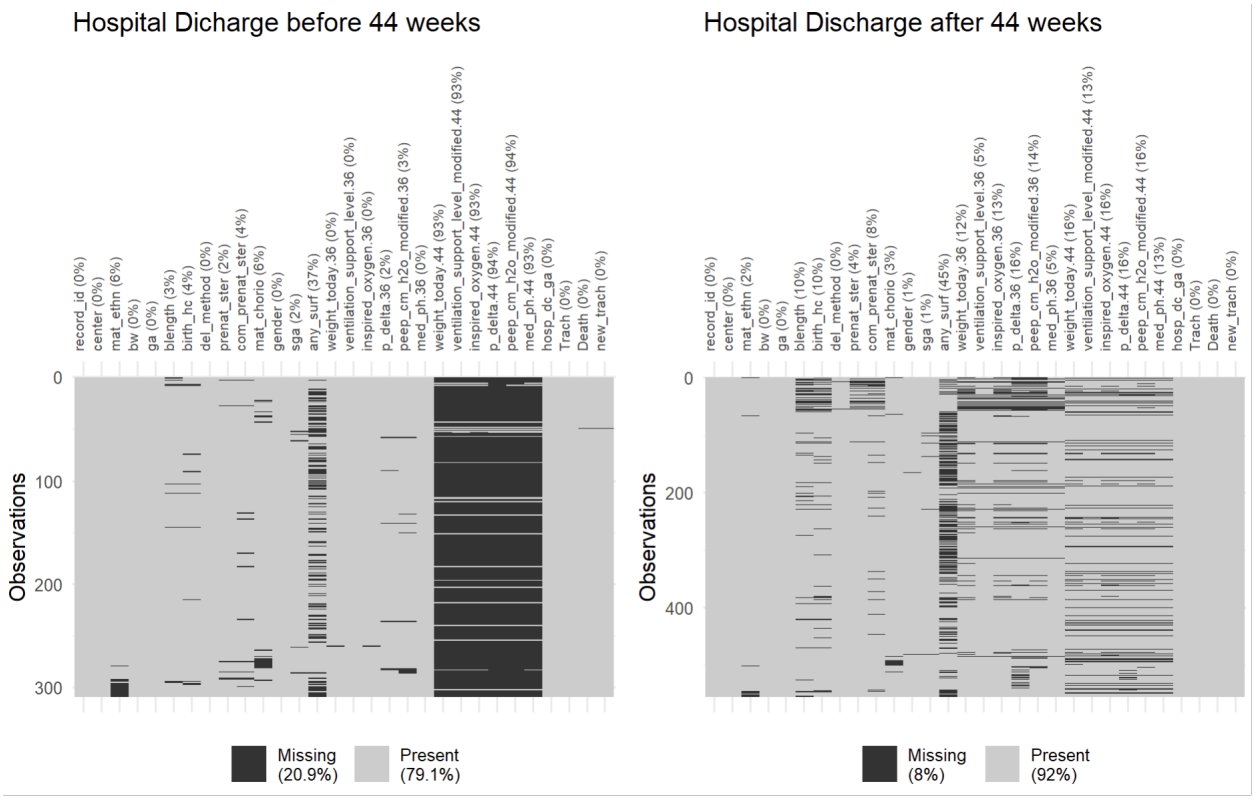

Figure 1: Missing Data

3 Results

The study employs three distinct models to discern predictive factors influencing patient outcomes across different stages of care. The Birth Model, encompassing the entire patient cohort, exclusively considers variables measured at birth. The 36-Week Model, focusing on patients with a hospital discharge occurring after at least 36 weeks, extends its analysis to include variables measured both at birth and at the 36-week mark. Lastly, the 44-Week Model, limited to patients with a hospital discharge after a minimum of 44 weeks, integrates all available variables throughout the entire care timeline.

3.1 Birth Model

The coefficients for the birth model are presented in Table 2. The mixed model employed in this analysis incorporates essential covariates, ensuring a comprehensive examination of factors influencing outcomes at birth. After applying the StepAIC criterion, the model considers gestational age, birth length, the use of prenatal steroids, and an indicator for infants categorized as small for gestational age. Each coefficient in the table signifies the change in the log-odds of the response variable associated with a one-unit change in the corresponding predictor variable, holding all other variables constant. Infants considered small for gestational age and infants whose mothers used prenatal steroids have positive coefficients implying they are more likely to have a tracheostomy. In the context of the LASSO model, it is noteworthy that the majority of coefficients exhibit smaller magnitudes compared to the coefficients from other models. However, we can emphasize that the signs of the coefficients align, enabling consistent conclusions to be drawn.

Table 3 presents the performance metrics for both the LASSO model and the comparative model. We can observe that the mixed model has higher AUC and sensitivity and a very close Brier score suggesting it performs better than the LASSO model. The recorded Sensitivity scores of 0.604 and 0.530, while indicative of reasonable performance, suggest a propensity for an increased likelihood of false negatives in the predictions. To enhance predictive accuracy, additional models at 36 weeks and 44 weeks have been incorporated, aiming to improve sensitivity and overall predictive capabilities. This strategic expansion across different time frames contributes to a comprehensive evaluation of predictive performance, refining the reliability of models in anticipating tracheotomies during the early stages of infant development.

	Mixed Model			**LASSO Model**		
Coefficient	Estimate	Coefficient	Estimate	Estimate	Coefficient	Estimate
intercept	-1.448	center2	-0.142	-0.270	center2	-1.316
mat_ethn2	0	center3	-1.061	0.0.354	center3	-2.096
bw	0	center4	0.550	0	center4	-0.475
ga	0.045	center5	-0.074	0.001	center5	-1.011
blength	-0.064	center7	-1.233	-0.0301	center7	-2.369
birth_hc	0	center12	2.005	-0.005	center12	0.843
del_method2	0	center16	-1.394	0.172	center16	-2.209
prenat_sterYes	0.568	center20	0.322	0.348	center 20	-0.195
com_prenat_sterYes	0	center1	1.700	0.036		
mat_chorioYes	0			0.0353		
genderMale	0			0.014		
sgaSGA	0.549			0.553		

Table 4: Coefficients for Birth Model

Metrics	Mixed Model	LASSO
AUC	0.750	0.745
Brier	0.129	0.128
Sensitivity	0.604	0.530
Specificity	0.813	0.873
Threshold	0.196	0.240

Table 5: Performance Metrics and Threshold for Predictions for the Generalized Mixed Model and LASSO Birth Model

3.2 36-Week Model

Table 5 presents the coefficients derived from both the mixed model, incorporating center as a random effect, and the lasso model, where center is treated as a fixed effect in the context of the 36-week models. Employing the StepAIC criterion in the mixed model resulted in the selection of predictors, including ethnicity, birthweight, birth head circumenference, prenatal steroids, weight at 36 weeks, type of ventilation support at 36 weeks, the fraction of inspired oxygen at 36 weeks, an indicator for medication for pulmonary hypertension at 36 weeks, positive and exploratory pressure at 36 week and the use of medication for pulmonary hypertension at 36 weeks. As the birth model, the presence of use of prenatal steroids emerge as a factor associated with an increased likelihood of tracheostomy. Also, a higher weight at 36 weeks is linked to a decreased probability of undergoing tracheostomy. Additionally, the log odds indicate that patients with non-invasive positive

pressure support exhibit lower odds compared to those with no support, while patients with invasive positive pressure support have higher odds. Furthermore, infants receiving medication for pulmonary hypertension at 36 weeks demonstrate elevated odds of tracheostomy. In the lasso model, birth weight and head circumference are not chosen as a significant variables even thought they were chosen in the StepAIC criterion. The consistent sign of coefficients indicates similar conclusions in the interpretation of coefficients between the two models.

Table 6 reveals an increase in performance metrics when predicting tracheostomy with data available at 36 weeks. The mixed model demonstrates a slight increase in AUC while having equal Brier score as the LASSO model, indicating improved discriminatory performance. Additionally, it is noteworthy that the Mixed Model exhibits higher sensitivity.

	Mixed Model			**LASSO Model**		
Coefficient	Estimate	Coefficient	Estimate	Estimate	Coefficient	Estimate
intercept	-3.817	center2	-3.200	-3.200	center2	-1.066
mat_ethn2	0.014	center3	-1.697	0.404	center3	-1.065
bw	0.001	center4	0.377	0.000	center4	-0.616
ga	0	center5	-0.034	0.004	center5	-2.742
blength	0	center7	-0.472	0.000	center7	-1.067
birth_hc	0.014	center12	1.773	0.017	center12	0.855
del_method2	0	center16	-0.867	0.021	center16	-1.707
prenat_sterYes	1.394	center20	-0.090	0.980	center20	-0.714
com_prenat_sterYes	0	center1	1.485	0.033		
mat_chorioYes	0			0.072		
genderMale	0			0.030		
sgaSGA	0			0.119		
weight_today.36	-0.001			-0.001		
ventilation_support_level.361	-0.890			-0.486		
ventilation_support_level.362	0.981			0.958		
inspired_oxygen.36	3.352			2.928		
p_delta.36	0			0.019		
peep_cm_h2o_modified.36	0.112			0.055		
med_ph.361	0.590			0.495		

Table 6: Coefficients for 36-Week Model

Metrics	Mixed Model	LASSO
AUC	0.905	0.903
Brier	0.094	0.094
Sensitivity	0.895	0.864
Specificity	0.800	0.807
Threshold	0.143	0.159

Table 7: Performance Metrics for the Generalized Mixed Model and LASSO 36-Week Model

3.3 44-Week Model

Table 7 presents coefficients of the mixed model with center as a random effect and LASSO with center as a fixed effect. Following the StepAIC criterion, the mixed model identified birth-weight, use of prenatal steroids, type of ventilation support at 36 weeks and 44 weeks, inspired oxygen at 36 weeks, peak inspiraroty pressure at 36 and 44 weeks, weight at 44 weeks, positive and exploratory pressure at 44 weeks and medication for pulmonary hypertension at 44 weeks. Notably, the significance of prenatal steroids extends across all three

models, while ventilation support and fraction of inspired oxygen at 36 weeks are also emphasized in the 36-week model. The model also highlights the importance of ventilation support at 44 weeks, indicating lower log odds for non-invasive positive pressure compared to no support and higher odds for invasive positive pressure compared to no support. Additionally, elevated positive and exploratory pressure at 44 weeks and medication for pulmonary hypertension are associated with increased odds of tracheostomy. The lasso model culminated in a sparser model when compared birth LASSO model and 36-week LASSO model. Still encompasses more variables than the StepAIC criterion. Notably, the lasso model excludes birth weight, gender, weight at 36 weeks, peek inspiratory pressure at 36 weeks, weight at 44 week, the first level of ventilation support (non-invasive positive pressure). This outcome underscores a limitation of LASSO, as it tends to eliminate specific levels of a variable rather than the variable as a whole. This highlights the need for careful consideration when interpreting and comparing variable selection outcomes between LASSO and alternative methods.

Table 8 illustrates the enhanced performance of this model in comparison to the birth and 36-week models. The mixed model exhibits higher AUC scores and lower Brier scores, indicative of improved discriminatory ability and overall model accuracy. Additionally, the mixed model demonstrates higher sensitivity and equal specificity. These findings contribute to an understanding of the model's strengths and trade-offs in predicting tracheostomy outcomes, emphasizing the importance of considering multiple performance metrics for a comprehensive evaluation.

	Mixed Model			**LASSO Model**		
Coefficient	Estimate	Coefficient	Estimate	Estimate	Coefficient	Estimate
intercept	-4.184	center2	-0.381	-4.077	center2	-0.306
mat_ethn2	0	center3	-0.584	0.050	center3	-0.366
bw	0.001	center4	-.9905	0.000	center4	-1.055
ga	0	center5	0.104	0.002	center5	0.000
blength	0	center7	-.803	0.002	center7	-0.662
birth_hc	0	center12	1.359	0.007	center12	1.292
del_method2	0	center16	0.514	0.076	center16	0.604
prenat_sterYes	1.866	center20	0.269	0.810	center20	0.216
com_prenat_sterYes	0	center1	0.784	0.135		
mat_chorioYes	0			0.028		
genderMale	0			0.000		
sgaSGA	0			0.013		
weight_today.36	0			0.000		
ventilation_support_level.361	-0.547			-0.356		
ventilation_support_level.362	-0.020			0.234		
inspired_oxygen.36	2.895			2.179		
p_delta.36	0.015			0.000		
peep_cm_h2o_modified.36	0			-0.002		
med_ph.361				0.054		
weight_today.44	-0.001			0.000		
ventilation_support_level_modified.441	-1.144			0.000		
ventilation_support_level_modified.442	0.610			1.302		
inspired_oxygen.44	0			0.260		
p_delta.44	-0.005			0.002		
peep_cm_h2o_modified.44	0.266			0.125		
med_ph.44	1.094			1.011		

Table 8: Coefficients for 44-Week Model

Metrics	Mixed Model	LASSO
AUC	0.938	0.904
Brier	0.085	0.086
Sensitivity	0.911	0.895
Specificity	0.864	0.864
Threshold	0.263	0.298

Table 9: Performance Metrics for the Generalized Mixed Model and LASSO 44-Week Model

4 Conslusion

4.1 Discussion

In this study, the objective was to construct predictive models at three critical time points in an infant's life: Birth, 36 weeks, and 44 weeks. A comparative analysis between two statistical methodologies—namely, a mixed model incorporating center as a random effect and a LASSO model with center as a fixed effect—was conducted. Given variable correlations, a StepAIC criterion was applied to the mixed model to facilitate variable selection, while the LASSO model inherently performs this task. Across all time point models, both methodologies exhibit commendable performance in key metrics. The utilization of prenatal steroids consistently proves significant across all six models. Notably, for predicting tracheostomy at birth, gestational age, birth length, and indicators for small gestational age were deemed as important variables. The 36-week mixed model underscores the significance of variables measured at this juncture, including weight, ventilation support, fraction of inspired oxygen, and an indicator for medication for pulmonary hypertension in predicting tracheostomy outcomes at 36 weeks. The 44-week mixed model is characterized as a sparse model, featuring ventilation support and a fraction of inspired oxygen at 36 weeks, ventilation support, positive and exploratory pressure, and medication for pulmonary hypertension at 44 weeks. Intriguingly, across all three-time points, the LASSO model does not exhibit the same degree of sparsity as the mixed model after applying the StepAIC criterion. This observation underscores that, despite the inherent variable selection capability of LASSO, the mixed model with StepAIC tends to yield a more sparsely parameterized model for this data finishing with a more interpretable model.

Given that AUC is a comprehensive metric assessing the overall discriminatory ability of a model, our model selection criterion prioritizes higher AUC scores. For all models, the AUC was always higher for the mixed model. Our recommendation leans toward a mixed model for prediction at each time point. This selection approach aligns with the distinctive strengths and performance characteristics demonstrated by each model variant at different time points in the predictive modeling process.

4.2 Limitations

Our work is subject to certain limitations. Notably, the dataset exhibits a substantial number of missing values, and the variable representing maternal race had to be excluded due to inconsistencies in the data. Furthermore, it is crucial to acknowledge the use of the StepAIC. It is important to recognize that stepAIC is implemented not primarily for the purpose of improving model performance, but rather to streamline model complexity while mitigating the impact on performance. In this context, alternative techniques, such as forward best subsets, could be explored. Additionally, an alternative approach involves fitting a mixed-effects model with a LASSO penalty, incorporating center as a random intercept term instead of treating

it as a fixed effect. These considerations underscore the need for a judicious exploration of alternative methodologies and the implications of specific choices in the modeling process.

Furthermore, the absence of transformations and interactions represents a significant limitation, primarily due to the potential of certain interactions to enhance the predictive capacity of the model. Hence, within the scope of limitations and future prospects, there exists the necessity to explore interactions among these variables for improved predictive modeling.

An important limitation observed in employing the LASSO model lay in its omission of randomization according to the center. Comparative evaluation against mixed models revealed superior predictive metrics within the mixed models when assessing the fixed effects of center in the LASSO model. This comparison vividly highlighted the unmistakable presence of a nested data structure, signaling a crucial need for its due consideration. This study acknowledges this limitation and presents it as a focal point for future research. An area of significant interest lies in the development and utilization of an adapted LASSO model incorporating random effects stratified by center. By addressing this, we aim to rectify the existing constraint inherent in the StepAIC criterion, while concurrently accommodating the nested structure inherent in the dataset. Such an approach is anticipated to offer an improved and more comprehensive framework for modeling, potentially enhancing the model's predictive capacity.

5 Code Apendix

```r
#Set working directoty
setwd("C:/Users/monic/OneDrive/Desktop/PHP2050/Project 2/")

#Libraries
library(readr)
library(naniar)
library(gtsummary)
library(tableone)
library(dplyr)
library(ggplot2)
library(gtable)
library(kableExtra)
library(mice)
library(corrplot)
library(ggpubr)
library(lme4)
library(MASS)
library(pROC)
library(glmnet)

#Import Data
df <- read.csv("project2.csv")


#Exclude duplicates ID
df <- df %>% filter(record_id != 2000824)



#Change classes
df <- df %>%
  mutate_if(is.character,as.factor)
df <- df %>%
  mutate(center = factor(center),
         mat_race = factor(mat_race),
         mat_ethn = factor(mat_ethn),
         del_method = factor(del_method),
         ventilation_support_level.36 = factor(ventilation_support_level
             .36),
         ventilation_support_level_modified.44 = factor(ventilation_
             support_level_modified.44),
```

```
         Trach = factor(Trach)
  )
df <- df %>%
  mutate_if(is.integer, as.numeric)
df <- df %>%
  mutate(Trach = case_when(Trach == "0" ~ "No",
                           Trach == "1" ~ "Yes")) %>%
  mutate(Trach = as.factor(Trach))

#which(df$center==21)
df$center[806] <- 20

df <- df %>% dplyr::select(!mat_race) #droppin race because of coding
    error
df <- df %>%
  mutate(com_prenat_ster = case_when(prenat_ster == "No" ~ "No", TRUE ~ as
      .character(com_prenat_ster)))
df <- df %>% mutate(com_prenat_ster = as.factor(com_prenat_ster))
df <- df %>% droplevels()
```

```
#Summary Statistics
tab_death <- df %>% dplyr::select(!record_id,-center) %>%
  tbl_summary(digits = list(everything() ~ c(2)),
                    statistic = list(all_continuous() ~ "{mean} ({sd})"),
                    by = Death,
              missing = "no") %>%
  add_overall() %>%
  add_n() %>%
  modify_header(label ~ "**Death**") %>%
  modify_spanning_header(c("stat_1", "stat_2") ~ "**Treatment Received**")
      %>%
  bold_labels()

tab_trach <- df %>% dplyr::select(!record_id,-center) %>%
  tbl_summary(digits = list(everything() ~ c(2)),
              statistic = list(all_continuous() ~ "{mean} ({sd})"),
              by = Trach,
              missing = "no") %>%
  add_overall() %>%
  add_n() %>%
  modify_header(label ~ "**Death**") %>%
  modify_spanning_header(c("stat_1", "stat_2") ~ "**Treatment Received**")
      %>%
```

```r
  bold_labels()

merged_table <- tbl_merge(
  tbls = list(tab_death, tab_trach),
  tab_spanner = c("**Death**", "**Tracheostomy**")) %>%
  as_kable_extra(booktabs = TRUE, caption = "Summary Statistics Grouped by Response Outcome") %>%
  kableExtra::kable_styling(latex_options = "scale_down")

merged_table

latex_code <- kable(merged_table, format = "latex", booktabs = TRUE, caption = "Summary Statistics Grouped by Response Outcome") %>%
  kable_styling(latex_options = "scale_down") %>%
  as.character()

# Print or save the LaTeX code
#cat(latex_code)
#writeLines(merged_table)

tab <- df %>% dplyr::select(center, Death, Trach, bw,weight_today.36, del_method, prenat_ster,ventilation_support_level.36,inspired_oxygen.36, p_delta.36, med_ph.36, inspired_oxygen.36) %>%
  tbl_summary(digits = list(everything() ~ c(2)),
                    statistic = list(all_continuous() ~ "{mean} ({sd})"),
                    by = center,
                missing = "no") %>%
  add_overall() %>%
  add_n() %>%
  modify_spanning_header( ~ "**Center**") %>%
  bold_labels() %>%
  as_kable_extra(booktabs = TRUE, caption = "Summary Statistics by Center") %>%
  kableExtra::kable_styling(latex_options = "scale_down")

latex_code <- kable(tab, format = "latex", booktabs = TRUE, caption = "Summary Statistics Grouped by Center") %>%
  kable_styling(latex_options = "scale_down") %>%
  as.character()

#cat(latex_code)

#save_as_image
```

```r
#library(webshot)
```

```r
#Composite Outcome
df <- df %>%
  mutate(new_trach = case_when( Trach == "Yes" ~ "Yes",
                          (Trach == "No" & Death == "Yes") ~ "Yes",
                          (Trach == "No" & Death == "No") ~ "No")
)
df$new_trach <- as.factor(df$new_trach)
#Remove any_surf
df <- df %>% dplyr::select(-any_surf)
```

```r
 #Creating three data frames for the three models
df_birthmodel <- df[,c(1:14,30)]
df_36wkmodel <- df[,c(1:20,27,30)]
df_44wkmodel <- df %>% dplyr::select(!Trach & !Death)

#Impute
#Birthmodel
df_birthmodel_out <- mice(df_birthmodel[,-1], 5, pri=F, seed = 1)
df_imp_birthmodel <- vector("list",5)
for (i in 1:5){df_imp_birthmodel[[i]] <- mice::complete(df_birthmodel_out,
   i)}


#36-wk model
df_36wk_out <- mice(df_36wkmodel[,-1], 5, pri=F, seed = 1)
df_imp_36wk <- vector("list",5)
for (i in 1:5){df_imp_36wk[[i]] <- mice::complete(df_36wk_out,i)}
#select correct population
for (i in 1:5) {
  df_imp_36wk[[i]] <- df_imp_36wk[[i]] %>% filter(hosp_dc_ga > 36) %>%
     dplyr::select(!hosp_dc_ga)
}

#44-wk model
df_44wk_out <- mice(df_44wkmodel[,-1], 5, pri=F, seed = 1)
df_imp_44wk <- vector("list",5)
for (i in 1:5){df_imp_44wk[[i]] <- mice::complete(df_44wk_out,i)}
#select correct population
for (i in 1:5) {
  df_imp_44wk[[i]] <- df_imp_44wk[[i]] %>% filter(hosp_dc_ga > 44) %>%
     dplyr::select(!hosp_dc_ga)
```

```
}
```

```
#pct_complete_case(df)


#Creating Missing Values Table
missing_table <- df %>% dplyr::select(!record_id) %>%
  summarize(across(everything(), ~ sum(is.na(.x)))) %>%
  t() %>%
  as.data.frame() %>%
  mutate(n=V1) %>%
  dplyr::select(n) %>%
  arrange(desc(n)) %>%
  mutate("%" = round(n/dim(df)[1],4)*100) %>%
  filter(n>0)


#kable(missing_table, caption = 'Missing Data per Variables', booktabs = TRUE) %>%
#  kable_styling(full_width=T, font_size = 8 ,latex_options = c('hold_position', 'scale_down'))

missing_table$Variable <- rownames(missing_table)
dd <- missing_table %>% dplyr::select(Variable, n, '%') #taking what we are going to separate
dd2 <- cbind(dd[1:13, ],dd[14:26,])#separating
kable(dd2,
      caption = "Missing Data per Variables",booktabs=T,row.names = F,
      align = "lrr") %>%
  kable_styling( font_size = 8,latex_options = c('scale_down'))


#Visualization for Missing Data
mis_vars <- missing_table %>% filter(n > 0) %>% rownames()

#vis_miss(df %>% select(all_of(mis_vars)))+theme(axis.text.x=element_text(size=rel(.9), angle = 90))
#vis_miss(df)+theme(axis.text.x=element_text(size=rel(.9), angle = 90))

missing_table_death_no <- df %>% filter(Death != "No") %>%
  summarize(across(everything(), ~ sum(is.na(.x)))) %>%
  t() %>%
  as.data.frame() %>%
```

```
  mutate(n=V1) %>%
  dplyr::select(n) %>%
  arrange(desc(n)) %>%
  mutate("%" = round(n/dim(df %>% filter(Death != "No"))[1],4)*100)
missing_table_death_yes <- df %>% filter(Death != "Yes") %>%
  summarize(across(everything(), ~ sum(is.na(.x)))) %>%
  t() %>%
  as.data.frame() %>%
  mutate(n=V1) %>%
  dplyr::select(n) %>%
  arrange(desc(n)) %>%
  mutate("%" = round(n/dim(df %>% filter(Death != "Yes"))[1],4)*100)



#vis_miss(df)
p1 <- vis_miss(df %>% filter(hosp_dc_ga < 44))+theme(axis.text.x=element_
    text(size=rel(.8), angle = 90))+ggtitle("Hospital Dicharge before 44
    weeks")
p2 <- vis_miss(df %>% filter(hosp_dc_ga > 44))+theme(axis.text.x=element_
    text(size=rel(.8), angle = 90))+ggtitle("Hospital Discharge after 44
    weeks")
ggarrange(p1,p2,ncol = 2)
```

```
##########Divide each imputed dataset into testing and training
#BIRTHMODEL
training_sets_birthmodel <- list()
testing_sets_birthmodel <- list()

set.seed(1)

for (i in 1:5) {
  n <- nrow(df_imp_birthmodel[[i]])
  num_test <- ceiling(n * .25)

  test_indices <- sample(1:n, num_test)

  training <- df_imp_birthmodel[[i]][-test_indices, ]
  testing <- df_imp_birthmodel[[i]][test_indices, ]

  training_sets_birthmodel[[i]] <- training
  testing_sets_birthmodel[[i]] <- testing
}
```

```r
#36-weekmodel
training_sets_36wk <- list()
testing_sets_36wk <- list()

set.seed(1)

for (i in 1:5) {
  n <- nrow(df_imp_36wk[[i]])
  num_test <- ceiling(n * .25)

  test_indices <- sample(1:n, num_test)

  training <- df_imp_36wk[[i]][-test_indices, ]
  testing <- df_imp_36wk[[i]][test_indices, ]

  training_sets_36wk[[i]] <- training
  testing_sets_36wk[[i]] <- testing
}

#44-weekmodel

training_sets_44wk <- list()
testing_sets_44wk <- list()

set.seed(1)

for (i in 1:5) {
  n <- nrow(df_imp_44wk[[i]])
  num_test <- ceiling(n * .25)

  test_indices <- sample(1:n, num_test)

  training <- df_imp_44wk[[i]][-test_indices, ]
  testing <- df_imp_44wk[[i]][test_indices, ]

  training_sets_44wk[[i]] <- training
  testing_sets_44wk[[i]] <- testing
}
```

```r
#Variable Selection for Mixed Model
birth_model_var <- vector("list",5)

for (i in 1:5) {
```

```r
xx <- training_sets_birthmodel[[i]][,-1] #not consider center(nested)
xx$new_trach <- as.factor(xx$new_trach)
model <- glm(new_trach ~., data = xx, family = binomial) %>%
  stepAIC(trace = FALSE)
ss <- summary(model)
ss$coefficients
birth_model_var[[i]] <- ss$coefficients
}


#Selected Variables for mixed model
#sapply(birth_model_var, function(df) rownames(df)[-1]) %>% table()
vars <- c(NA)
for (i in 1:5) {
  vars <- c(vars,unlist(rownames(birth_model_var[[i]])))
}
vars <- table(vars[-1])
vars <- vars[vars>=2]
vars <- names(vars)

coef_birthmodel_mm <- matrix(NA,nrow = length(vars), ncol = 5)
rownames(coef_birthmodel_mm) <- vars
coef_center__birthmodel_mm <-  matrix(NA,nrow = 9, ncol = 5)
rownames(coef_center__birthmodel_mm) <- c("center1","center2","center3","
    center4","center5","center7","center12","center16","center20")
for (i in 1:5) {
  training_sets_birthmodel[[i]]$new_trach <- as.factor(training_sets_
      birthmodel[[i]]$new_trach)
  mod <- glmer(new_trach ~ blength + ga + prenat_ster + sga +(1 | center),
               data = training_sets_birthmodel[[i]] ,
               family = binomial,
               control=glmerControl(optimizer="bobyqa",
                                    optCtrl=list(maxfun=2e5)))
  k <- mod %>% summary() %>% coef()
  coef_birthmodel_mm[,i] <- k[,1]
  k <- ranef(mod)$center
  coef_center__birthmodel_mm[,i] <- k$`(Intercept)`

}

coef_birthmodel_mm <- rowMeans(coef_birthmodel_mm)
coef_center__birthmodel_mm <- rowMeans(coef_center__birthmodel_mm)

```

```r

##Predict
auc_birthmodel_mm <- rep(NA,5)
brier_birthmodel_mm <- rep(NA,5)
sensitivity_birthmodel_mm <- rep(NA,5)
specificity_birthmodel_mm <- rep(NA,5)
threshold_birthmodel_mm <- rep(NA,5)

for (i in 1:5) {
x_new <- testing_sets_birthmodel[[i]]
x_new <- x_new %>% dplyr::select(blength,com_prenat_ster,ga, prenat_ster, sga, new_trach,center)
x_new <- cbind(1,x_new)
x_new <- model.matrix(new_trach ~ blength+ ga+ prenat_ster+sga + center, x_new,
                      contrasts.arg = list(center = contrasts( testing_sets_birthmodel[[i]]$center,
                                                               contrasts = FALSE)))

pred <- x_new %*% c(coef_birthmodel_mm,coef_center__birthmodel_mm)
pred <- exp(pred)/(exp(pred)+1)

levels(testing_sets_birthmodel[[i]]$new_trach) = c(0,1)

roc1 <- roc(predictor = pred, response = testing_sets_birthmodel[[i]]$new_trach, levels = c(0,1), direction = "<")
#plot(roc1, print.thres=TRUE)

k <- coords(roc=roc1, x = "best")

#store metrics
auc_birthmodel_mm[i] <- auc(roc1)
brier_birthmodel_mm[i] <- mean((pred - (as.numeric(testing_sets_birthmodel[[i]]$new_trach)-1))^2)
sensitivity_birthmodel_mm[i] <- k$sensitivity
specificity_birthmodel_mm[i] <- k$specificity
threshold_birthmodel_mm[i] <- k$threshold
}

auc_birthmodel_mm_pooled <- mean(auc_birthmodel_mm)
brier_birthmodel_mm_pooled <- mean(brier_birthmodel_mm)
sensitivity_birthmodel_mm_pooled <- mean(sensitivity_birthmodel_mm)
```

```r
specificity_birthmodel_mm_pooled <- mean(specificity_birthmodel_mm)
threshold_birthmodel_mm_pooled <- mean(threshold_birthmodel_mm)
```

```r
#LASSO
lasso <- function(df) {
  #' Runs 10-fold CV for lasso and returns corresponding coefficients
  #' @param df, data set
  #' @return coef, coefficients for minimum cv error

  # Matrix form for ordered variables
  x.ord <- model.matrix(new_trach~., data = df)[,-1]
  y.ord <- df$new_trach

  k <- 10
  set.seed(1)
  folds <- sample(1:k, nrow(df), replace=TRUE)

  # Lasso model
  lasso_mod_cv <- cv.glmnet(x.ord, y.ord, nfolds = 10, foldid = folds,
                            alpha = 1, family = "binomial")
  lasso_mod <- glmnet(x.ord, y.ord, nfolds = 10, foldid = folds,
                            alpha = 1, family = "binomial",
                            lambda = lasso_mod_cv$lambda.min)
  # Get coefficients
  coef <- coef(lasso_mod)
  return(coef)
}

# Find average lasso coefficients over imputed datasets
lasso_coef1 <- lasso(training_sets_birthmodel[[1]])
lasso_coef2 <- lasso(training_sets_birthmodel[[2]])
lasso_coef3 <- lasso(training_sets_birthmodel[[3]])
lasso_coef4 <- lasso(training_sets_birthmodel[[4]])
lasso_coef5 <- lasso(training_sets_birthmodel[[5]])
lasso_coef <- cbind(lasso_coef1, lasso_coef2, lasso_coef3,
                    lasso_coef4, lasso_coef5)
avg_coefs_lasso_birthmodel <- apply(lasso_coef, 1, mean)


df_birthmodel_long_test <- bind_rows(testing_sets_birthmodel)

x_vars <- model.matrix(new_trach ~. , df_birthmodel_long_test) #get long
    data model matrix
```

```r

pred <- x_vars %*% avg_coefs_lasso_birthmodel #predict
pred <- exp(pred)/(1+exp(pred))


y <- df_birthmodel_long_test$new_trach
levels(y) = c(0,1)
roc1 <- roc(predictor = pred, response = y, levels = c(0,1), direction = "<")
#plot(roc1, print.thres=TRUE)
k <- coords(roc=roc1, x = "best")

#store metrics
auc_birthmodel_lasso <- auc(roc1)[1]
brier_birthmodel_lasso <- mean((pred - (as.numeric(y)-1))^2)
sensitivity_birthmodel_lasso <- k$sensitivity
specificity_birthmodel_lasso <- k$specificity
threshold_birthmodel_lasso <- k$threshold
```

```r
##################36wk model
#Variable Selection for Mixed Model
wk36_var <- vector("list",5)

for (i in 1:5) {
xx <- training_sets_36wk[[i]][,-1] #not consider center(nested)
model <- glm(new_trach ~., data = xx, family = binomial) %>%
  stepAIC(trace = FALSE)
ss <- summary(model)
ss$coefficients
wk36_var[[i]] <- ss$coefficients
}


#Selected Variables for mixed model
#sapply(birth_model_var, function(df) rownames(df)[-1]) %>% table()
vars <- c(NA)
for (i in 1:5) {
  vars <- c(vars,unlist(rownames(wk36_var[[i]])))
}
vars <- table(vars[-1])
vars <- vars[vars>=2]
vars <- names(vars)

```

```r
coef_36wk_mm <- matrix(NA,nrow = length(vars), ncol = 5)
rownames(coef_36wk_mm) <- vars
coef_center__36wk_mm <-  matrix(NA,nrow = 9, ncol = 5)
rownames(coef_center__36wk_mm) <- c("center1","center2","center3","center4","center5","center7","center12","center16","center20")
for (i in 1:5) {
  mod <- glmer(new_trach ~ birth_hc + bw + inspired_oxygen.36 + med_ph.36 + peep_cm_h2o_modified.36 + prenat_ster +
                ventilation_support_level.36 + weight_today.36 + (1 | center),
              data = training_sets_36wk[[i]] ,
              family = binomial,
              control=glmerControl(optimizer="bobyqa",
                                   optCtrl=list(maxfun=2e5)))
  k <- mod %>% summary() %>% coef()
  coef_36wk_mm[,i] <- k[,1]
  k <- ranef(mod)$center
  coef_center__36wk_mm[,i] <- k$`(Intercept)`
  }

coef_36wk_mm <- rowMeans(coef_36wk_mm)
coef_center__36wk_mm <- rowMeans(coef_center__36wk_mm)


##Predict
auc_36wk_mm <- rep(NA,5)
brier_36wk_mm <- rep(NA,5)
sensitivity_36wk_mm <- rep(NA,5)
specificity_36wk_mm <- rep(NA,5)
threshold_36wk_mm <- rep(NA,5)

for (i in 1:5) {
x_new <- testing_sets_36wk[[i]]
x_new <- x_new %>% dplyr::select( new_trach, birth_hc, bw, inspired_oxygen.36,  med_ph.36, peep_cm_h2o_modified.36, prenat_ster ,
                ventilation_support_level.36 , weight_today.36, center)
#x_new <- cbind(1,x_new)
x_new <- model.matrix(new_trach ~   birth_hc+ bw+ inspired_oxygen.36+  med_ph.36+peep_cm_h2o_modified.36+prenat_ster +
                ventilation_support_level.36 + weight_today.36+ center, x_new,
                    contrasts.arg = list(center = contrasts( testing_sets_36wk[[i]]$center,
```

```
                                                                    contrasts =
                                                                      FALSE)))

pred <- x_new %*% c(coef_36wk_mm,coef_center__36wk_mm)
pred <- exp(pred)/(exp(pred)+1)

levels(testing_sets_36wk[[i]]$new_trach) = c(0,1)

roc1 <- roc(predictor = pred, response = testing_sets_36wk[[i]]$new_trach,
    levels = c(0,1), direction = "<")
#plot(roc1, print.thres=TRUE)

k <- coords(roc=roc1, x = "best")

#store metrics
auc_36wk_mm[i] <- auc(roc1)
brier_36wk_mm[i] <- mean((pred - (as.numeric(testing_sets_36wk[[i]]$new_
   trach)-1))^2)
sensitivity_36wk_mm[i] <- k$sensitivity
specificity_36wk_mm[i] <- k$specificity
threshold_36wk_mm[i] <- k$threshold
}

auc_36wk_mm_pooled <- mean(auc_36wk_mm)
brier_36wk_mm_pooled <- mean(brier_36wk_mm)
sensitivity_36wk_mm_pooled <- mean(sensitivity_36wk_mm)
specificity_36wk_mm_pooled <- mean(specificity_36wk_mm)
threshold_36wk_mm_pooled <- mean(threshold_36wk_mm)
```

```
# Find average lasso coefficients over imputed datasets
lasso_coef1 <- lasso(training_sets_36wk[[1]])
lasso_coef2 <- lasso(training_sets_36wk[[2]])
lasso_coef3 <- lasso(training_sets_36wk[[3]])
lasso_coef4 <- lasso(training_sets_36wk[[4]])
lasso_coef5 <- lasso(training_sets_36wk[[5]])
lasso_coef <- cbind(lasso_coef1, lasso_coef2, lasso_coef3,
                    lasso_coef4, lasso_coef5)
avg_coefs_lasso_36wk <- apply(lasso_coef, 1, mean)


df_36wk_long_test <- bind_rows(testing_sets_36wk)

x_vars <- model.matrix(new_trach ~. , df_36wk_long_test) #get long data
```

```
    model matrix

pred <- x_vars %*% avg_coefs_lasso_36wk #predict
pred <- exp(pred)/(1+exp(pred))


y <- df_36wk_long_test$new_trach
levels(y) = c(0,1)
roc1 <- roc(predictor = pred, response = y, levels = c(0,1), direction = "
    <")
#plot(roc1, print.thres=TRUE)
k <- coords(roc=roc1, x = "best")

#store metrics
auc_36wk_lasso <- auc(roc1)[1]
brier_36wk_lasso <- mean((pred - (as.numeric(y)-1))^2)
sensitivity_36wk_lasso <- k$sensitivity
specificity_36wk_lasso <- k$specificity
threshold_36wk_lasso <- k$threshold
```

```
##################44wk model
#Variable Selection for Mixed Model
wk44_var <- vector("list",5)

for (i in 1:5) {
xx <- training_sets_44wk[[i]][,-1] #not consider center(nested)
model <- glm(new_trach ~., data = xx, family = binomial) %>%
  stepAIC(trace = FALSE)
ss <- summary(model)
ss$coefficients
wk44_var[[i]] <- ss$coefficients
}


#Selected Variables for mixed model
#sapply(birth_model_var, function(df) rownames(df)[-1]) %>% table()
vars <- c(NA)
for (i in 1:5) {
  vars <- c(vars,unlist(rownames(wk44_var[[i]])))
}
vars <- table(vars[-1])
vars <- vars[vars>=2]
vars <- names(vars)
```

```r

coef_44wk_mm <- matrix(NA,nrow = length(vars), ncol = 5)
rownames(coef_44wk_mm) <- vars
coef_center__44wk_mm <-  matrix(NA,nrow = 9, ncol = 5)
rownames(coef_center__44wk_mm) <- c("center1","center2","center3","center4","center5","center7","center12","center16","center20")
for (i in 1:5) {
  mod <- glmer(new_trach ~ bw+ inspired_oxygen.36 + med_ph.44 + p_delta.36+
                 p_delta.44 + peep_cm_h2o_modified.44+
                 prenat_ster + ventilation_support_level.36 + ventilation_support_level_modified.44 + weight_today.44 +(1 | center),
               data = training_sets_44wk[[i]] ,
               family = binomial,
               control=glmerControl(optimizer="bobyqa",
                                    optCtrl=list(maxfun=2e5)))
  k <- mod %>% summary() %>% coef()
  coef_44wk_mm[,i] <- k[,1]
  k <- ranef(mod)$center
  coef_center__44wk_mm[,i] <- k$`(Intercept)`

}

coef_44wk_mm <- rowMeans(coef_44wk_mm)
coef_center__44wk_mm <- rowMeans(coef_center__44wk_mm)


##Predict
auc_44wk_mm <- rep(NA,5)
brier_44wk_mm <- rep(NA,5)
sensitivity_44wk_mm <- rep(NA,5)
specificity_44wk_mm <- rep(NA,5)
threshold_44wk_mm <- rep(NA,5)

for (i in 1:5) {
x_new <- testing_sets_44wk[[i]]
x_new <- x_new %>% dplyr::select( bw, inspired_oxygen.36, med_ph.44 , p_delta.36,              p_delta.44 ,peep_cm_h2o_modified.44,
                 prenat_ster ,ventilation_support_level.36 , ventilation_support_level_modified.44 , weight_today.44 , new_trach, center)
x_new <- model.matrix(new_trach ~  bw+inspired_oxygen.36+ med_ph.44+ p_
```

```
   delta.36+         p_delta.44+ peep_cm_h2o_modified.44+
                  prenat_ster +ventilation_support_level.36 + ventilation_
                     support_level_modified.44 + weight_today.44 +center, x
                     _new,
                         contrasts.arg = list(center = contrasts( testing_
                            sets_36wk[[i]]$center,
                                                                  contrasts =
                                                                     FALSE)))

pred <- x_new %*% c(coef_44wk_mm,coef_center__44wk_mm)
pred <- exp(pred)/(exp(pred)+1)

levels(testing_sets_44wk[[i]]$new_trach) = c(0,1)

roc1 <- roc(predictor = pred, response = testing_sets_44wk[[i]]$new_trach,
    levels = c(0,1), direction = "<")
#plot(roc1, print.thres=TRUE)

k <- coords(roc=roc1, x = "best")

#store metrics
auc_44wk_mm[i] <- auc(roc1)
brier_44wk_mm[i] <- mean((pred - (as.numeric(testing_sets_44wk[[i]]$new_
   trach)-1))^2)
sensitivity_44wk_mm[i] <- k$sensitivity
specificity_44wk_mm[i] <- k$specificity
threshold_44wk_mm[i] <- k$threshold

}

auc_44wk_mm_pooled <- mean(auc_44wk_mm)
brier_44wk_mm_pooled <- mean(brier_44wk_mm)
sensitivity_44wk_mm_pooled <- mean(sensitivity_44wk_mm)
specificity_44wk_mm_pooled <- mean(specificity_44wk_mm)
threshold_44wk_mm_pooled <- mean(threshold_44wk_mm)
```

```
# Find average lasso coefficients over imputed datasets
lasso_coef1 <- lasso(training_sets_44wk[[1]])
lasso_coef2 <- lasso(training_sets_44wk[[2]])
lasso_coef3 <- lasso(training_sets_44wk[[3]])
lasso_coef4 <- lasso(training_sets_44wk[[4]])
lasso_coef5 <- lasso(training_sets_44wk[[5]])
lasso_coef <- cbind(lasso_coef1, lasso_coef2, lasso_coef3,
```

```
                    lasso_coef4, lasso_coef5)
avg_coefs_lasso_44wk <- apply(lasso_coef, 1, mean)


df_44wk_long_test <- bind_rows(testing_sets_44wk)

x_vars <- model.matrix(new_trach ~. , df_44wk_long_test) #get long data
    model matrix

pred <- x_vars %*% avg_coefs_lasso_44wk #predict
pred <- exp(pred)/(1+exp(pred))


y <- df_44wk_long_test$new_trach
levels(y) = c(0,1)
roc1 <- roc(predictor = pred, response = y, levels = c(0,1), direction = "
    <")
#plot(roc1, print.thres=TRUE)
k <- coords(roc=roc1, x = "best")

#store metrics
auc_44wk_lasso <- auc(roc1)[1]
brier_44wk_lasso <- mean((pred - (as.numeric(y)-1))^2)
sensitivity_44wk_lasso <- k$sensitivity
specificity_44wk_lasso <- k$specificity
threshold_44wk_lasso <- k$threshold
```

```
res_metrics <- data.frame(Metrics = c("AUC", "BRIER", "Sensitivity", "
    Specificity", "Threshold")
)
res_birthmodel <- data.frame(Mixed_Model = c(auc_birthmodel_mm_pooled,
    brier_birthmodel_mm_pooled,sensitivity_birthmodel_mm_pooled,specificity
    _birthmodel_mm_pooled,threshold_birthmodel_mm_pooled),
            LASSO = c(auc_birthmodel_lasso,brier_birthmodel_lasso,
                sensitivity_birthmodel_lasso,specificity_birthmodel_lasso,
                threshold_birthmodel_lasso))

rownames(res_birthmodel) <- c("AUC", "BRIER", "Sensitivity", "Specificity"
    , "Threshold")


res_36wk <- data.frame(Mixed_Model = c(auc_36wk_mm_pooled,brier_36wk_mm_
    pooled,sensitivity_36wk_mm_pooled,specificity_36wk_mm_pooled,threshold_
```

```
    36wk_mm_pooled),
10          LASSO = c(auc_36wk_lasso,brier_36wk_lasso,sensitivity_36wk_
                lasso,specificity_36wk_lasso,threshold_36wk_lasso))
11
12 rownames(res_36wk) <- c("AUC", "BRIER", "Sensitivity", "Specificity", "
    Threshold")
13
14
15 res_44wk <- data.frame(Mixed_Model = c(auc_44wk_mm_pooled,brier_44wk_mm_
    pooled,sensitivity_44wk_mm_pooled,specificity_44wk_mm_pooled,threshold_
    44wk_mm_pooled),
16          LASSO = c(auc_36wk_lasso,brier_44wk_lasso,sensitivity_44wk_
                lasso,specificity_44wk_lasso,threshold_44wk_lasso))
17
18 rownames(res_44wk) <- c("AUC", "BRIER", "Sensitivity", "Specificity", "
    Threshold")
19
20
21 #kable(list(res_metrics, res_birthmodel, res_36wk, res_44wk), booktabs = T
    , row.names = F, caption = "Model Results") %>%
22 # add_header_above(c(" ", "Birth Model", "36-Week Model", "44-Week Model")
    ) %>%
23 #  kable_styling(latex_options = c('scale_down'))
```