Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ ___4/14/2025___

Date

Investigation of the association between preterm infants' red blood cell transfusion risk and maternal-neonatal characteristics

By

Xintian Song
MSPH

Department of Biostatistics and Bioinformatics

_____

Amita Manatunga (Committee Chair)

_____

Kirk Easley (Committee Member)

Investigation of the association between preterm infants' red blood cell transfusion needs and maternal-neonatal characteristics

By

Xintian Song

B.S., Guangdong Technion – Israel Institute of Technology, 2018

Thesis Committee Chair: Amita Manatunga, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the
degree of Master of Science in Public Health in
Department of Biostatistics and Bioinformatics
2025

Abstract

Investigation of the association between preterm infants' red blood cell transfusion risk and maternal-neonatal characteristics

By Xintian Song

Objectives: Red blood cell (RBC) transfusions are a crucial intervention for preterm infants, with the need for transfusions varying based on maternal and neonatal factors. Understanding these factors can enhance clinical decision-making and optimize transfusion practices.

Method: This study investigates the association between preterm infants' RBC transfusion needs and key maternal and neonatal characteristics using three modeling approaches: (1) Quasi-Poisson regression to estimate transfusion rates, (2) the Anderson-Gill model to assess transfusion intensity over time, and (3) the Frailty model to account for individual heterogeneity in relative transfusion risk.

Results: This study highlights the significant influence of maternal and neonatal characteristics on the RBC transfusion needs of preterm infants. Among the three models applied, the multivariable Frailty model provided the most robust risk estimates by accounting for individual heterogeneity. The NICU center where an infant was born played a crucial role in transfusion outcomes. While a higher proportion of infants born at Grady received at least one transfusion compared to those at Emory and Northside, the Frailty model revealed that the adjusted relative risk of transfusion was lower at Grady (RR = 0.88, 95% CI: 0.58, 1.34). After adjusting for center, maternal race, and maternal age, each additional week of gestational age reduced the relative risk of transfusion by 15% (RR = 0.85; 95% CI: 0.78, 0.92, $P < .001$). Similarly, every 100g increase in birth weight was associated with a 23% reduction in transfusion risk (RR = 0.77; 95% CI: 0.71, 0.83, $P < .001$). Additionally, the adjusted RR of transfusion per 1g/dL decrease in hemoglobin at birth was 1.11; 95% CI: 1.04, 1.17, $P < .001$).

Conclusion: The estimated risk of transfusions for all covariates are similar across three models while the Frailty model indicating a moderate within-subject correlation between transfusion times. Despite relatively homogenous preterm infant population, gestational age, birth weight, and hemoglobin levels at birth remained significant predictors of transfusion risk. Furthermore, we found the transfusion risks are significantly different across centers.

Investigation of the association between preterm infants' red blood cell transfusion risk and maternal-neonatal characteristics

By

Xintian Song

B.S., Guangdong Technion – Israel Institute of Technology, 2018

Thesis Committee Chair: Amita Manatunga, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the
degree of Master of Science in Public Health in
Department of Biostatistics and Bioinformatics
2025

Table of Contents

1. Introduction 1

2. Method 4

1.1 Study Population 4

1.2 Statistical Analysis 5

1.2.1 Quasi Poisson Model 6

1.2.2 Anderson Gill Model 7

1.2.3 Frailty Model 8

3. Results 10

3.1 Description of Exposures and Outcomes 10

3.2 Risk Factors for RBC Transfusion Needs 15

3.2.1 Quasi Poisson Model 15

3.2.2 Anderson Gill Model 16

3.2.3 Frailty Model 17

3.3 Risk Factors for Low Birth Weight 22

4. Discussion 24

5. References 29

1. Introduction

Premature birth, anemia, birth complications, and neonatal infections are among the leading causes of infant mortality (World Health Organization, 2024). Preterm infants, those born before 37 completed weeks of gestation, often have low birth weight and require specialized medical care. Red blood cell (RBC) transfusions are commonly administered in this population to improve oxygen delivery and manage anemia ("Guidelines for Transfusion of Pediatric Patients," 2016). Given the high vulnerability of preterm infants, understanding their postnatal clinical needs, such as RBC transfusions, is critical to improving survival outcomes.

However, RBC transfusions are not without risks. One major concern is necrotizing enterocolitis (NEC), a serious gastrointestinal disease. NEC is a leading cause of death in preterm infants, particularly between 2 weeks and 2 months of age, and carries a fatality rate of 20% to 30% in neonatal intensive care units (Patel et al., 2015). Receipt of RBC transfusions has been linked to an increased risk of NEC. Patel et al. (2016) employed a multivariable competing-risks Cox regression model and found that the risk of NEC was significantly higher among very low birth weight infants with severe anemia compared to those without severe anemia at any given week [HR = 5.99 (95% CI, 2.00-18.0); P = .001]. To mitigate these risks, it is essential to study neonates most in need of RBC transfusions.

Few studies have examined the clinical characteristics of very low birth weight infants that require RBC transfusions. Hosono et al. (2006) identified significant differences in gestational age, birth weight, baseline hemoglobin levels, Apgar scores, and phlebotomy

loss between infants who received at least one transfusion and those who did not. The logistic regression analysis highlighted key predictive factors for transfusion, including lower baseline hemoglobin levels (OR = 2.61 per 1 g/dL decrease), lower birthweight (OR = 3.00 per 100 g decrease), and smaller gestational age (OR = 1.89 per week decrease). Similarly, Ekhaguere et al. (2016) conducted a multivariable logistic regression analysis and reported that the likelihood of receiving at least one RBC transfusion decreased with higher baseline hemoglobin level [OR = 0.70 (95% CI, 0.65–0.75) per 1 g/dL increase], female sex [OR = 0.64 (95% CI, 0.47–0.87)], and higher birth weight [OR = 0.71 (95% CI, 0.65–0.76) per 100 g increase].

Infants with low birth weight represent a significant subpopulation of concern in neonatal care. Adugna and Worku (2022) highlighted several maternal characteristics associated with an increased risk of low birth weight. Key factors include preterm birth [OR = 38.0 (95% CI, 15.3–93.0)], pregnancy-induced hypertension [OR = 2.6 (95% CI, 1.1–6.4)], and a maternal body mass index (BMI) of < 18.5 kg/m^2 [OR = 6.8 (95% CI, 1.5–31.1)], all of which significantly increase the likelihood of delivering an infant with low birth weight. Xi et al. (2020), in their exploration of maternal lifestyle behaviors, found that several factors were significantly associated with low birth weight in preterm infants compared to the control group. These characteristics included maternal age over 35 years [OR = 2.49 (95% CI, 1.10–5.76)], gestational hypertension [OR = 9.06 (95% CI, 3.70–22.33)], and exposure to passive smoking [OR = 1.40 (95% CI, 1.05–1.85)].

Among studies examining the association between red blood cell transfusion needs and various maternal and infant characteristics, most do not account for multiple transfusions

within a given infant. In this study, we focus on the outcome as repeated blood transfusions per infant. Since the multiple transfusion events within an infant are likely to be correlated, statistical methods that account for within-subject dependence are required. Additionally, some infants may leave the study early due to discharge or other factors, leading to right censoring (Therneau and Grambsch 2000). To address the challenges, we apply a series of approaches tailored to the recurrent nature of transfusion events, appropriately interpreting results within the assumptions of each model.

The objectives of this thesis are as followed. The first is to further investigate the mechanism of how maternal and infant characteristics contribute to frequent red blood cell transfusions. Motivated by the process, we would also want to investigate how neonatal characteristics like infant birthweight can serve as a mediator on the causal pathway between maternal characteristics and transfusion outcomes. Additionally, we would like to explore a series of statistical methods on how to model repeated transfusion outcomes while handling censoring, followed by a discussion on the interpretation of the results.

2. Method

1.1 Study Population

The data were collected from a longitudinal, prospective, multicenter observational cohort study at the 3 hospitals in Atlanta, GA (Grady Memorial Hospital, Emory University Hospital Midtown, and Northside Hospital (Marin T et al., 2018). Per protocol 220 very low preterm infants were enrolled a (recruited from 2017 to 2021 with a pause due to COVID-19). The cohort enrollment as of June 2024 was 284 infants. There were no standardized transfusion practice guidelines in place for these 3 institutions, allowing for an evaluation of different transfusion strategies and varying thresholds for initiating treatment, all while ensuring adherence to the standards of care (Patel et al., 2016).

To maintain the integrity of the cohort, specific exclusion criteria were applied. Infants were excluded if they were not expected to survive beyond seven days based on the neonatologist's assessment, had severe congenital abnormalities, received a transfusion prior to enrollment, or if their mothers chose not to participate (Patel et al., 2016).

Given the nature of the study, each infant in the cohort might undergo multiple transfusion events over the 90-day follow-up period or until hospital discharge. However, some infants exited the study earlier due to reasons such as, mother declined consent, death from complications, etc. Despite these variations in follow-up duration, all RBC transfusion exposures were meticulously documented by research nurses using case report forms.

1.2 Statistical Analysis

Descriptive statistics were summarized for subjects who received at least one RBC transfusion and those who did not. Categorical variables were presented as proportions, while continuous variables were described using mean and standard deviation. A multivariate linear regression was conducted to examine the association between birth weight and key neonatal and maternal characteristics.

Since each infant in the cohort might experience multiple transfusions, repeated transfusion outcomes were recorded based on the start and end time of each event episode. Three approaches were employed to model multiple blood transfusion events: (1) The rate of RBC transfusions was estimated using a Quasi-Poisson regression; (2) Transfusion intensity, considering event time, was evaluated with the Anderson-Gill model; and (3) To account for individual variability, the relative risk of RBC transfusions was estimated using the Frailty model, an extension of the Cox proportional hazards model.

The major difference among them is the way the repeated events are modelled. The Poisson model assumes that transfusion counts are independent among infants. In contrast, the Anderson-Gill model assumes a Markov process, where future events depend only on the immediate past. The Frailty model, however, accounts for dependency among the recurrent event times by introducing shared subjects' random effects (Amorim and Cai 2014).

All models were adjusted for potential confounders like center. The selection of confounders was based on prior knowledge on neonatal care and appropriate model selections. All the analysis were done by using glm function from R base and coxph function from R survival

package. Different options on the coxph are specified to fit Anderson-Gill and Frailty model respectively.

1.2.1 Quasi Poisson Model

The Poisson model generalizes transfusion events as a counting process over the study follow-up period. The outcome of interest represents the average transfusion rate within the population. It assumes an underlying Poisson distribution for transfusion counts with a scale parameter ϕ to account for overdispersion. This model is appropriate when the timing of recurrent events is not a primary focus, but when an understanding of the overall rate is desired.

The risk rate of the i_{th} subject is modeled as:

$$\lambda_i = E(Y_i \mid X_i) = \exp(\beta X_i), i = 1, \cdots \cdots, n,$$

where λ_i refers to the expected rate of transfusion occurrences for subject i given covariates X_i, Y_i refers to the transfusion counts, $\exp(\beta)$ refers to the rate ratio corresponding to covariates X_i and is assumed to be constant over time.

The variance of transfusion counts is model as:

$$Var(Y_i \mid X_i) = \phi\lambda_i,$$

where ϕ refers to the overdispersion parameter and is assumed to be greater than 1.

The coefficients of β is estimated by likelihood function:

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(-\lambda)\cdot\lambda^{y_i}}{y_i!} = \prod_{i=1}^{n} \frac{e^{-\exp(\beta X_i)}\cdot\exp(\beta X_i)^{y_i}}{y_i!}$$

1.2.2 Anderson Gill Model

The Anderson-Gill model extends the proportional hazard Cox model by focusing on increments in the number of events over time. Similar to the Generalized Estimating Equations (GEE) approach, a robust sandwich estimator is used to account for the correlation of recurrent events within the same individual. The model assumes the baseline risk h_0 is shared across recurrent events. It is of notice that Anderson-Gill model does not consider the order of recurrences, and the dependency is only captured by observed covariates. Thus, its primary goal in this study is to estimate the overall effect, rather than event-specific effects, by examining the intensity of transfusion occurrences.

The risk intensity for the i_{th} subject is modeled as:

$$h_i(t_{ij} \mid X_i) = h_0(t_{ij}) \exp(\beta X_i), i = 1, \cdots\cdots, n, and\ j = 1, \cdots\cdots, k_i, and\ k_i \leq k,$$

where h_i refers to the risk of subject i given covariates X_i, h_0 refers to the baseline risk for all the subjects, $\exp(\beta)$ refers to the risk ratio given covariates X_i and is assumed to be constant over time.

The coefficients of β is estimated by partial likelihood function:

$$L(\beta) = \prod_{i=1}^{n} \prod_{j=1}^{k_i} \frac{\exp(\beta X_i)}{\sum_{l \in R_{AG}} \exp(\beta X_i)},$$

where the risk set of the model is given by:

$$R_{AG}(t_{ij}) = \{i: \exists j \in \{1, \ldots\ldots, k_i, such\ that\ t_{ij} \geq t\},$$

which means that prior to event time t, those subjects i who neither censored nor having not experienced their last events k_i remains into the risk set (Ozga, Kieser, and Rauch 2018).

1.2.3 Frailty Model

The Frailty Model treats frailties – the excess risk for different subjects, as an unobserved covariate. The Frailty Model assumes that the observed hazard is influenced by both covariates and unmeasured individual-level correlation. The outcome of interest is person-specific relative risk.

The relative risk for the i_{th} subject is modeled by:

$$h_i(t_{ij} \mid X_i, \omega_i) = h_0(t_{ij})\, \omega_i \exp(\beta X_i)$$

where h_i refers to the risk of subject i given covariates X_i and frailty ω_i, ω_i is assumed to remain constant for all relevant event j for i_{th} subject. Frailty ω_i could has a log-Gamma or log-normal distribution, with $\exp(\omega_i) \sim G(\frac{1}{\theta}, \frac{1}{\theta})$ or $\exp(\omega_i) \sim N(0, \theta)$ respectively. The unknown parameter θ is a dispersion parameter.

The baseline hazard is different for each subject given by $h_0(t_{ij})\omega_i$, such that each subject has their own disposition to failure. Subjects have an elevated risk of the event compared to the average will have $\omega_i > 1$. If θ= 0, there is no frailty effect, and the model reduces to a standard Cox model. Thus, within the subject i, one unit of increase in covariates X_i could lead to $\exp(\beta)$ in the relative risk.

The coefficients of β is estimated by partial likelihood function:

$$L(\beta) = \prod_{i=1}^{n} \prod_{j=1}^{k_i} \int_0^{\infty} \frac{\omega_i \exp(\beta X_i)}{\sum_{I \in R_{Frailty}} \omega_i \exp(\beta X_i)} \cdot f(\ \omega_i \mid \theta)\ d\omega_i,$$

where the risk set of the model is given by:

$$R_{Frailty}(t_{ij}) = \{i: \exists j \in \{1, \ldots \ldots, k_i, such\ that\ t_{ij} \geq t\},$$

which means that prior to event time t, those subjects i who neither censored nor having not experienced their last events k_i remains into the risk set (Ozga, Kieser, and Rauch 2018).

Model selection between Gamma or Gaussian frailty term can be guided by criteria such as the Restricted Maximum Likelihood (REML) method, which balances model complexity and goodness-of-fit.

2. Results

2.1 Description of Exposures and Outcomes

The study population consisted of 284 preterm infants, among whom 109 did not receive any transfusions during the study period, there are 11 out of 109 infants had no recorded follow-up visits and were excluded from the analysis if censoring incorporated.

Across three NICUs, the proportion of whether infants receiving transfusions ever varied (Table 1). 59% and 57% of infants received a transfusion in Emory and Northside Hospital respectively, compared to 76% at Grady. These differences suggest that the likelihood of receiving a transfusion for preterm infants may depend on the NICUs they were born to.

Preterm infants who received at least one transfusion were compared to those who did not across various maternal characteristics. The median maternal age was 31 years for mothers of preterm infants who received transfusion ever, with a similar interquartile range to those whose infants did not receive a transfusion. A greater proportion of African American mothers had infants who received at least one transfusion (64%) compared to non-Black mothers (57%). Infants born to mothers with lower educational had a higher likelihood of receiving RBC transfusions, as 83% of mothers with education before college delivered infants who had transfusion, compared to 44% among those mothers with education beyond college. Infants having transfusion needs were more common in those mothers who have substance use during pregnancy (72% vs. 60%). Mode of delivery showed a minor difference, with 60% of infants delivered via Caesarean section receiving a transfusion later

compared to 68% of those delivered by other methods. Regarding steroid usage during delivery, infants whose mothers received no steroid treatment contribute to higher transfusion needs (74%), while those mothers prescribed with Betamethasone had a lower transfusion needs (62%). The presence of prenatal complications and the number of fetuses did not show substantial variation in transfusion needs (Table 2).

Additionally, neonatal characteristics were compared between preterm infants who underwent at least one transfusion and those who did not. Infants who never received a transfusion had a significantly higher mean gestational age (28.7 weeks) compared to those who underwent at least one transfusion (26.4 weeks). Similarly, the non-transfused group had a higher mean birth weight (1,051g) than the transfused group (824.8g). Hemoglobin levels at birth were also notably higher among infants who were never transfused (17.2g/dL) compared to those who ever received transfusions (14.0 g/dL). The average SNAP score, which assesses illness severity in newborns, where a higher score indicates greater mortality risk, was slightly elevated in the transfused group (6.7) compared to those who were never transfused (6.0). However, both scores fell within the mild category, suggesting the cohort was generally expected to have favorable future survival outcomes (Maiya, Nagashree, and Shaik, 2001). Furthermore, the mean Apgar score at 1 minute, which reflects how well the infant tolerated the birthing process, was higher among infants who received transfusions (5.8) compared to those who did not (4.3). The mean Apgar score at 5 minutes, assessing the infant's condition after birth, was also higher in the transfused group (3.0) compared to the non-transfused group (2.2) (Table 3). Both groups had Apgar scores significantly below 7, indicating the cohort's need for intensive medical attention after delivery. However, these

scores did not necessarily indicate infants' future health outcomes ("Apgar Score: MedlinePlus Medical Encyclopedia").

Table 1: Enrollment characteristics of NICUs in relation to whether preterm infants received transfusion ever or not

		Whether Infants Received Transfusion?		
	Overall N = 284[1]	No (n=109)	Yes (n=177)	**p-value**[1]
Center				**0.026**
Emory	51	21 (41%)	30 (59%)	
Grady	66	16 (24%)	50 (76%)	
Northside	167	72 (43%)	95 (57%)	

[1]P-value is calculated based on Pearson's Chi-squared test

Table 2: Descriptive characteristics of mothers in relation to whether their infants received transfusion ever or not

		Whether Infants Received Transfusion?		
	Overall N = 284	No (n=109)	Yes (n=177)	**p-value**[1]
Mother age (per 1 year old increases)				0.092
Mean (SD)	30 (6)	31 (6)	30 (6)	
Mother Race				0.4
Black	190	69 (36%)	121 (64%)	
Not Black	94	40 (43%)	54 (57%)	
Education[2]				**<0.001**
After College	50	28 (56%)	22 (44%)	
Before College	76	13 (17%)	63 (83%)	
Substance Abuse				0.3
Any use	32	9 (28%)	23 (72%)	
No use	251	100 (40%)	151 (60%)	
Final Delivery Mode				0.3
Caesarean section	226	90 (40%)	136 (60%)	
Other	57	18 (32%)	39 (68%)	
Steroid Type				**0.005**
Betamethasone	230	88 (38%)	142 (62%)	
No Steroid	38	10 (26%)	28 (74%)	
Other Steroid	13	10 (77%)	3 (23%)	
Prenatal Complication				0.4
Hypertension	58	27 (47%)	31 (53%)	
Multiple complications	16	6 (38%)	10 (63%)	
No complication	185	64 (35%)	121 (65%)	
Other complication	18	7 (39%)	11 (61%)	
Fetus Number[3]				0.7
>=2	79	28 (35%)	51 (65%)	
1	204	80 (39%)	124 (61%)	

[1]P-value is calculated based on Pearson's Chi-squared test for categorical variables and Welch Two Sample t-test for continuous variables

[2]Mothers' education level was unavailable for 158 subjects. The category 'After College' includes those who obtained a college degree or a higher-level graduate degree.

[3]Fetus Number refers to a singleton pregnancy vs. multiple gestation pregnancy

Table 3: Descriptive characteristics of infants in relation to whether they received transfusion ever or not

		Whether Infants Received Transfusion?		
	Overall N = 284	No (n=109)	Yes (n=177)	**p-value**[1]
Gestational Age (week)				**<0.001**
Mean (SD)	27.3 (2.4)	28.7 (2.2)	26.4 (2.2)	
Infant Birth weight (g)				**<0.001**
Mean (SD)	911.7 (221.8)	1,051.0 (175.4)	824.8 (202.9)	
Hemoglobin level at Birth (1g/dL) [2]				**<0.001**
Mean (SD)	15.2 (2.9)	17.2 (2.5)	14.0 (2.5)	
Days of breast milk feeding in 1st 10 days[3]				**0.029**
Mean (SD)	9.9 (0.5)	10.0 (0.0)	9.9 (0.6)	
SNAP score at Birth [4]				**<0.001**
Mean (SD)	6.4 (1.3)	6.0 (1.4)	6.7 (1.2)	
Apgar Score 1 min [5]				**<0.001**
Mean (SD)	5.2 (2.5)	4.3 (2.5)	5.8 (2.4)	
Apgar Score 5 min [6]				**0.002**
Mean (SD)	2.7 (2.1)	2.2 (1.9)	3.0 (2.1)	

[1]P-value is calculated based on Welch Two Sample t-test

[2]Hemoglobin level at Birth was unavailable for 35 subjects. It is estimated using the value from day 1, 2, or 3 after birth if the initial measurement is missing.

[3]Days of breast milk feeding in 1st 10 days was unavailable for 57 subjects

[4]SNAP score measure the illness severity for the newborns and was unavailable for 11 subjects

[5]The 1-minute score determines how well the infant tolerated the birthing process

[6]The 5-minute score tells how well the infant is doing outside the mother's womb

2.2 Risk Factors for RBC Transfusion Needs

2.2.1 Quasi Poisson Model

The univariate association suggests that compared to NICU at Emory Hospital, the Grady Hospital had a 17% reduction in transfusion rate [RR = 0.83, (95% CI, 0.52 - 1.35)], while the Northside Hospital showed a 47% reduction [RR = 0.53, (95% CI, 0.35 - 0.82)]. Infants delivered by Black mothers had a 62% higher transfusion intensity compared to non-Black mothers [RR = 1.62, (95% CI, 1.11 - 2.41)]. Gestational age and infant birth weight were significantly associated with transfusion rates. For every 1-week increase in gestational age, the transfusion rate decreased by 28% [RR = 0.72, (95% CI, 0.67 - 0.78)], and for each 100g increase in infant birth weight, the rate decreased by 30% [RR = 0.70, (95% CI, 0.66 - 0.75)]. Additionally, hemoglobin levels at birth showed a positive association, with a 25% increase in the transfusion rate for every 1g/dL decrease in hemoglobin [RR = 1.25, (95% CI, 1.18 - 1.33)]. Overall, the results suggest that delivered by Black mother, lower gestational age, lower birth weight, and lower hemoglobin levels at birth are associated with higher transfusion rate ratio (Table 4).

Adjusted for NICUs, maternal race, and maternal age, gestational age had a significant effect on transfusion rates, with each one-week increase associated with a 13% decrease in the transfusion rate ratio (95% CI: 0.79, 0.94, $P < .001$). Infant birth weight also showed a significant association, with each 100g increase in birth weight resulting in a 19% decrease in the transfusion rate ratio (95% CI: 0.74, 0.88, $P < .001$). Hemoglobin level at birth was

another significant factor, with each 1g/dL decrease in hemoglobin level associated with a 25% increase in the transfusion rate ratio (95% CI: 1.18, 1.33, P = 0.005) (Table 5).

2.2.2 Anderson Gill Model

The univariate association suggests that compared to NICU at Emory Hospital, Grady Hospital had a 10% reduction in transfusion intensity [RR = 0.90, (95% CI, 0.58 - 1.39)], and Northside Hospital had a 43% reduction [RR = 0.57, (95% CI, 0.38 - 0.85)]. Infants delivered by Black mothers had a 57% higher transfusion intensity compared to non-Black mothers [RR = 1.57, (95% CI, 1.08 - 2.28)]. Additionally, gestational age and birth weight remained significantly associated with transfusion intensity. Each 1-week increase in gestational age was associated with a 43% reduction in transfusion intensity [RR = 0.57, (95% CI, 0.38 - 0.85)], while each 100g increase in birth weight reduced the intensity by 32% [RR = 0.68, (95% CI, 0.64 - 0.72)]. Hemoglobin levels also contributed significantly to transfusion intensity, with a 26% increase for every 1g/dL decrease in hemoglobin [RR = 1.26, (95% CI, 1.20 - 1.34)]. Overall, the results suggest that delivered by Black mother, lower gestational age, lower birth weight, and lower hemoglobin levels at birth are associated with higher transfusion intensity ratio (Table 4).

Adjusted for NICUs, maternal race, and maternal age, gestational age remained significantly associated with transfusion intensity, where each one-week increase in gestational age resulted in a 15% decrease in the transfusion intensity ratio (95% CI: 0.77, 0.93, P < .001). Infant birth weight continued to show a significant inverse relationship with transfusion intensity, with each 100g increase in birth weight leading to a 21% reduction in transfusion

intensity (95% CI: 0.73, 0.86, P < .001). Additionally, each 1g/dL decrease in hemoglobin at birth was associated with a 26% increase in the transfusion intensity ratio (95% CI: 1.03, 1.13, P = 0.002) (Table 5).

2.2.3 Frailty Model

Frailty Model with a gamma distributed frailty is chosen to conduct the analyses. The univariate association suggests that compared to NICU at Emory Hospital, Grady Hospital had a 12% reduction in transfusion relative risk [RR = 0.88, (95% CI, 0.49 - 1.58)], and Northside Hospital had a 46% reduction [RR = 0.54, (95% CI, 0.33 - 0.88)]. Infants delivered by Black mothers had a 65% higher relative risk of transfusion compared to non-Black mothers [RR = 1.65, (95% CI, 1.13, 2.41)]. Gestational age, birth weight, and hemoglobin levels remained key factors in transfusion risk. A 1-week increase in gestational age decreased the relative risk by 29% [RR = 0.71, (95% CI, 0.67 - 0.76)], and each 100g increase in birth weight decreased the relative risk by 35% [RR = 0.65, (95% CI, 0.61 - 0.69)]. Hemoglobin levels at birth showed a strong association, with a 31% increase in relative risk for every 1g/dL decrease in hemoglobin [RR = 1.31, (95% CI, 1.24 - 1.39)]. Overall, the results suggest that delivered by Black mother, lower gestational age, lower birth weight, and lower hemoglobin levels at birth are associated with higher transfusion relative risk (Table 4).

Adjusted for NICUs, maternal race, and maternal age, gestational age again showed a significant effect on transfusion relative risk, with each one-week increase in gestational age resulting in a 15% reduction in the relative risk of transfusion (95% CI: 0.78, 0.92, P < .001). For infant birth weight, a 100g increase was associated with a 23% decrease in relative risk

(95% CI: 0.71, 0.83, P < .001). Hemoglobin level at birth also had a significant effect, with each 1g/dL decrease in hemoglobin level leading to a 11% increase in relative risk (95% CI: 1.04, 1.17, P < .001) (Table 5).

Additionally, the estimated random effect $\hat{\theta}$ from the Frailty model is 0.45, indicating a moderate within-subject correlation. As a result, the effect sizes derived from the Frailty model are similar to those obtained from other two models using robust standard error. However, accounting for within-subject variability does impact statistical significance. Specifically, when frailty is not considered, the effects of certain covariates may appear less statistically significant (Figure 1).

Table 4: Univariate association between transfusion needs and maternal and infant characteristics

	Transfusion Rate Ratio[1] (95% CI)	p-value	Transfusion Intensity Ratio[2] (95% CI)	p-value	Transfusion Relative Risk[3] (95% CI)	p-value
Center		**0.006**		**0.008**		**<0.001**
Emory	—		—		—	
Grady	0.83 (0.52, 1.35)		0.90 (0.58, 1.39)		0.88 (0.49, 1.58)	
Northside	0.53 (0.35, 0.82)		0.57 (0.38, 0.85)		0.54 (0.33, 0.88)	
Mother Race		**0.011**		**0.017**		**0.01**
Not Black	—		—		—	
Black	1.62 (1.11, 2.41)		1.57 (1.08, 2.28)		1.65 (1.13, 2.41)	
Mother Age[4]		>0.9		>0.9		0.79
Normal	—		—		—	
High	0.98 (0.66, 1.42)		0.99 (0.68, 1.44)		0.95 (0.63, 1.42)	
Gestational Age (per 1 week increase)	0.72 (0.67, 0.78)	**<0.001**	0.71 (0.66, 0.76)	**<0.001**	0.71 (0.67, 0.76)	**<0.001**
Infant Birth Weight (per 100g increase)	0.70 (0.66, 0.75)	**<0.001**	0.68 (0.64, 0.72)	**<0.001**	0.65 (0.61, 0.69)	**<0.001**
Hemoglobin Level at Birth (per 1g/dL decrease)	1.25 (1.18, 1.33)	**0.005**	1.26 (1.20, 1.34)	**<0.001**	1.31 (1.24, 1.39)	**<0.001**

[1]Derived from Quasi Poisson model

[2]Derived from Anderson-Gill model

[3]Derived from Frailty model

[4]High mother age is defined by higher than 35 years old

Table 5: Multivariate association between transfusion needs and maternal and infant characteristics

	Transfusion Rate Ratio[1] (95% CI)	p-value	Transfusion Intensity Ratio[2] (95% CI)	p-value	Transfusion Relative Risk[3] (95% CI)	p-value
Center		**<0.001**		**<0.001**		**<0.001**
Emory	—		—		—	
Grady	1.07 (0.73, 1.56)		1.12 (0.79, 1.57)		0.88 (0.58, 1.34)	
Northside	0.56 (0.40, 0.80)		0.57 (0.39, 0.83)		0.44 (0.30, 0.64)	
Mother Race		0.5		0.5		0.384
Not Black	—		—		—	
Black	0.90 (0.66, 1.25)		0.90 (0.66, 1.21)		0.87 (0.64, 1.19)	
Mother age[4]		0.2		0.095		0.182
Normal	—		—		—	
High	1.25 (0.92, 1.67)		1.26 (0.96, 1.66)		1.23 (0.91, 1.65)	
Gestational Age (per 1 week increase)	0.87 (0.79, 0.94)	**<0.001**	0.85 (0.77, 0.93)	**<0.001**	0.85 (0.78, 0.92)	**<0.001**
Infant Birth Weight (per 100g increase)	0.81 (0.74, 0.88)	**<0.001**	0.79 (0.73, 0.86)	**<0.001**	0.77 (0.71, 0.83)	**<0.001**
Hemoglobin Level at Birth (per 1g/dL decrease)	1.09 (1.03, 1.15)	**0.005**	1.08 (1.03, 1.13)	**0.002**	1.11 (1.04, 1.17)	**<0.001**

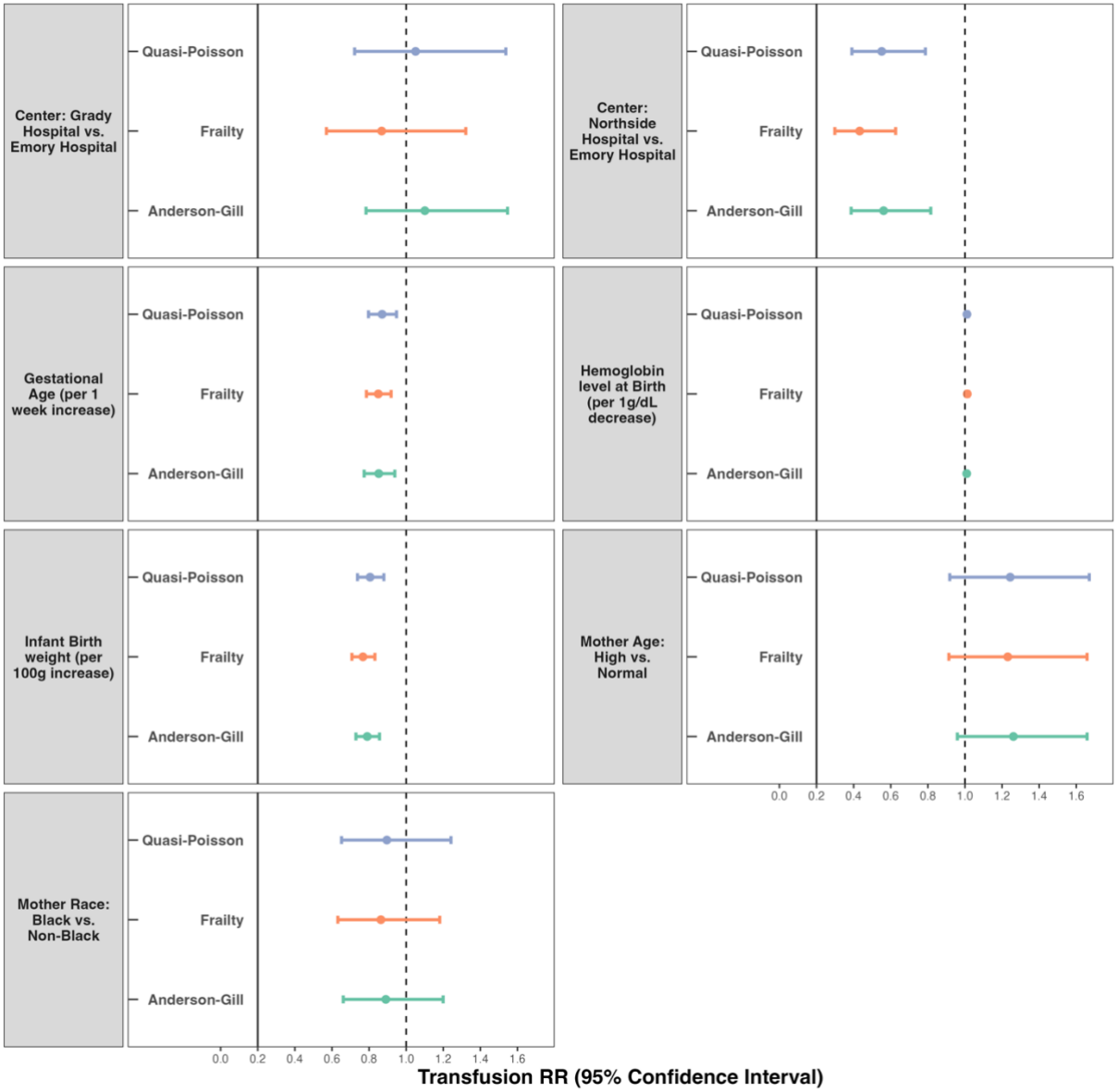
[1]Derived from Quasi Poisson model

[2]Derived from Anderson-Gill model

[3]Derived from Frailty model

[4]High mother age is defined by higher than 35 years old

Figure 1: Comparison for multivariate association between transfusion needs and maternal and infant characteristics



2.3 Risk Factors for Low Birth Weight

In the univariate analysis, the coefficient for center showed that infants born at Grady had a birth weight increase of 57 grams (95% CI: -25, 138, P = 0.4), and those born at Northside had an increase of 27 grams (95% CI: -43, 97, P = 0.4) compared to preterm infants born at Emory's NICU. For maternal race, infants born to Black mothers had a coefficient of -54 grams (95% CI: -109, 1.1, P = 0.055), indicating a slight decrease in birth weight compared to infants born to non-Black mothers. Maternal age was not significantly associated with birth weight, with a mean difference of -14 grams (95% CI: -79, 43, P = 0.6) for mothers who delivered at age 35 or older compared to those who delivered before 35.Gestational age had a significant positive effect, with each one-week increase in gestational age associated with an increase of 63 grams in birth weight (95% CI: 55, 71, P < .001).

In the multivariate analysis, after adjusting for NICUs, maternal race, and maternal age, gestational age remained significant, with each one-week increase still associated with an average of 65-gram increase in preterm infants' birth weight (95% CI: 57, 74, P < .001). 2 fetuses per gestation was associated with a 59-gram increase in preterm infants' birth weight compared with singleton, and 3 fetuses associated with a 52-gram decrease. Although it may not be entirely intuitive for multiple gestations to be associated with increased birth weight compared to singletons, the number of fetuses per gestation was identified a statistically significant factor influencing birth weight.

Table 6: The association between infants' birthweight and maternal and infant characteristics

	Coefficient[1]	95% CI	p-value	Coefficient[2]	95% CI	p-value
(Intercept)	—	—	—	-829	-1171, -664	**<0.001**
Center			0.4			0.22
Emory	—	—		—	—	
Grady	57	-25, 138		43	-16, 102	
Northside	27	-43, 97		44	-9.1, 97	
Mother Race			0.055			0.47
Not Black	—	—		—	—	
Black	-54	-109, 1.1		-16	-60, 28	
Mother Age[3]			0.6			0.082
Normal	—	—		—	—	
High	-14	-79, 43		-37	-79, 4.7	
Fetus Number[4]			**0.023**			**0.014**
1						
2	74	14, 135		59	15, 103	
3	-81	-228, 67		-52	-158, 54	
Gestational Age (per 1 week increase)	63	55, 71	**<0.001**	65	57, 74	**<0.001**

[1] Derived from univariate association

[2]Derived from multivariate association

[3]High mother age is defined by higher than 35 years old

[4]Fetus Number refers to a singleton pregnancy vs. multiple gestation pregnancy

3. Discussion

To investigate how maternal and neonatal characteristics contribute to multiple RBC transfusions, this thesis leverages multiple modeling approaches and longitudinal follow-up to assess transfusion risk with regard to multiple baseline mother and infant characteristics. Significant findings were observed after adjusting for center, maternal race, and maternal age, showing that infants born with shorter gestational age, lower birth weight, and lower hemoglobin levels were associated with higher transfusion needs. Among the three models, the Frailty model provided the most accurate estimates by accounting for within-subject dependence, yielding more robust inferences. Specifically, it showed that a one-week increase in gestational age, a 100-gram increase in birth weight, and a 1 g/dL decrease in hemoglobin levels at birth were associated with a 15%, 23%, and 11% change in relative transfusion risk, respectively.

Additionally, the NICU center where an infant is born plays a crucial role in transfusion outcomes and should be carefully adjusted. Preterm infants born at Grady's NICU have a higher proportion of receiving at least one transfusion compared to those born at Emory and Northside, yet Grady also have the highest average birth weight among the three centers. Without adjusting for subject heterogeneity, infants born at Grady appear to have the highest transfusion needs. However, after accounting for within-subject dependence, the relative transfusion risk is lower than at Emory, suggesting that the transfusion needs at Grady having a higher variability, which might be influenced by unobserved factors, such as

heterogeneous clinical practice standards for RBC transfusion, postnatal care protocols, and maternal population profiles.

However, there are several limitations that need further examination. One of them concerns data quality, particularly the incomplete collection of maternal socio-demographic factors such as maternal education level. Collection of this variable was discontinued midway through the study because it was not available in the medical records for all participants. Additionally, the Anderson-Gill and Frailty models used in this thesis assume a constant baseline hazard, which may fail to capture the impact of covariates whose effects change over time, particularly between the first transfusion event and subsequent recurrences. This limitation could lead to an underestimation of the covariate effect sizes (Amorim and Cai, 2014). Another limitation is that neither the Anderson-Gill nor the Frailty model explicitly accounts for the order of transfusion events, such that a subject is not at risk for the k_{th} event until the first $k-1_{th}$ events have occurred (Ozga et al., 2018). This could limit the ability to capture potential differences between early and later transfusion events.

To address these limitations, future work should prioritize improving data collection procedures to reduce missingness, particularly for key maternal socio-demographic variables. Moreover, incorporating time-varying effects, such as estimating differential risks $\exp(\beta_s)$ across distinct periods before, during, and after the NEC infection window (2 weeks to 60 days), may offer a more accurate reflection of changing covariate effect size over time. In parallel, exploring alternative recurrent event models, such as the Prentice, Williams, and Peterson (PWP) model, could enhance the analysis. As a stratified extension of the Anderson-Gill model, the PWP model introduces event-specific baseline hazards h_{0j} and

applies a more restricted risk set, allowing for better characterization of the event order and capturing covariate effects $\exp(\beta_j)$ specific to each transfusion event (Castañeda and Gerritse, 2010).

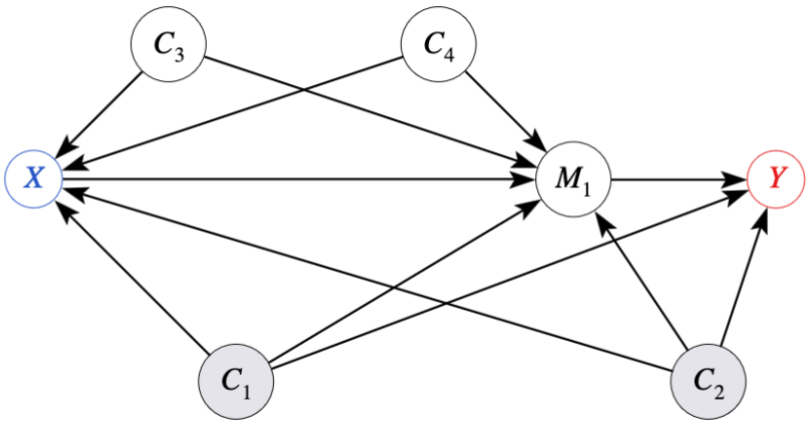
From the previous results, we conclude that higher birth weight is associated with reduced RBC transfusion needs among preterm infants, emphasizing the importance of promoting adequate weight gain during the gestational period. Given these findings, birthweight could serve as a mediator in the relationship between maternal race and repeated transfusion outcomes, after adjusting for relevant confounders such as maternal age and center. Additionally, factors that may confound the relationship between maternal race and birth weight, such as the number of fetuses per gestation and gestational age, should be accounted for to improve the robustness of the analysis (Figure 2).

To better understand the causal relationship between maternal race and transfusion needs, two key causal effects could be explored. First, estimating the conditional average treatment effect (CATE) would help identify subpopulations where maternal race has the greatest impact on transfusion rates, providing insight into how racial differences in transfusion needs vary across different subgroups of preterm infants. Second, assessing the natural indirect effect (NIE) would allow for the quantification of how much the effect of maternal race on transfusion outcomes is mediated through birth weight rather than direct effects of maternal race.

To sum up, including mediation analysis in the future analysis would provide deeper insights into the causal pathways linking maternal characteristics, neonatal factors, and transfusion

risk, helping to better understand racial disparities in transfusion needs. One might find what is the counterfactual birthweight under different racial group and the transfusion risk ratios between preterm infants of Black and non-Black mothers, while holding birth weight constant at the average level within each racial group.

Figure 2: Proposed DAG for causal pathway between maternal race and repeated transfusion needs



X - Maternal Race

Y – Repeated Transfusion Outcome

M_1 – Infant Birthweight

C_1 – Center

C_2 – Maternal Age at the Time of Delivery

C_3 – Gestational Age

C_4 – Fetus Number

4. References

1. Adugna, Dagnew Getnet, and Misganaw Gebrie Worku. 2022. “Maternal and Neonatal Factors Associated With Low Birth Weight Among Neonates Delivered at the University of Gondar Comprehensive Specialized Hospital, Northwest Ethiopia.” *Frontiers in Pediatrics* 10 (August). https://doi.org/10.3389/fped.2022.899922.
2. Amorim, Leila Daf, and Jianwen Cai. 2014a. “Modelling Recurrent Events: A Tutorial for Analysis in Epidemiology.” *International Journal of Epidemiology* 44 (1): 324–33. https://doi.org/10.1093/ije/dyu222.
3. “Apgar Score: MedlinePlus Medical Encyclopedia.” n.d. https://medlineplus.gov/ency/article/003402.html.
4. Castañeda, Javier, Bart Gerritse, 2010. “Appraisal of Several Methods to Model Time to Multiple Events per Subject: Modelling Time to Hospitalizations and Death.” Revista Colombiana De Estadística. Vol. 33.
5. Ekhaguere, Osayame A., Frank H. Morriss, Edward F. Bell, Nadkarni Prakash, and John A. Widness. 2016. “Predictive Factors and Practice Trends in Red Blood Cell Transfusions for Very-low-birth-weight Infants.” *Pediatric Research* 79 (5): 736–41. https://doi.org/10.1038/pr.2016.4.
6. “GUIDELINES FOR TRANSFUSION OF PEDIATRIC PATIENTS.” 2016. https://www.wadsworth.org/sites/default/files/WebDoc/ped_tx_guidelines_2.pdf.
7. Hosono, Shigeharu, Hideo Mugishima, Masami Shimada, Michiyoshi Minato, Tomo Okada, Shigeru Takahashi, and Kensuke Harada. 2006. “Prediction of Transfusions in

Extremely Low-birthweight Infants in the Erythropoietin Era." *Pediatrics International* 48 (6): 572–76. https://doi.org/10.1111/j.1442-200x.2006.02279.x.

8. Maiya, P. P., S. Nagashree, and Mehabood Shariff Shaik. 2001. "Role of Score for Neonatal Acute Physiology (SNAP) in Predicting Neonatal Mortality." *The Indian Journal of Pediatrics* 68 (9): 829–34. https://doi.org/10.1007/bf02762105.
9. Marin T, Patel RM, Roback JD, Stowell SR, Guo Y, Easley KA, Warnock M, Skvarich J and Josephson CD. Does red blood cell irradiation and/or anemia trigger intestinal injury in premature infants with birth weight ≤ 1250 grams? An observational birth cohort study. BMC Pediatrics 2018 Aug 11;18(1):270. doi: 10.1186/s12887-018-1241-5.
10. Ozga, Ann-Kathrin, Meinhard Kieser, and Geraldine Rauch. 2018. "A Systematic Comparison of Recurrent Event Models for Application to Composite Endpoints." BMC Medical Research Methodology 18 (1). https://doi.org/10.1186/s12874-017-0462-x.
11. Patel, R. M., Kandefer, S., Walsh, M. C., Bell, E. F., Carlo, W. A., Laptook, A. R., Sánchez, P. J., Shankaran, S., Van Meurs, K. P., Ball, M. B., Hale, E. C., Newman, N. S., Das, A., Higgins, R. D., & Stoll, B. J. (2015). Causes and Timing of Death in Extremely Premature Infants from 2000 through 2011. New England Journal of Medicine, 372(4), 331–340. https://doi.org/10.1056/nejmoa1403489
12. Patel, Ravi M., Andrea Knezevic, Neeta Shenvi, Michael Hinkes, Sarah Keene, John D. Roback, Kirk A. Easley, and Cassandra D. Josephson. 2016. "Association of Red Blood Cell Transfusion, Anemia, and Necrotizing Enterocolitis in Very Low-Birth-Weight Infants." *JAMA* 315 (9): 889. https://doi.org/10.1001/jama.2016.1204.

13. Thenmozhi, Mani, Visalakshi Jeyaseelan, Lakshmanan Jeyaseelan, Rita Isaac, and Rupa Vedantam. 2019. “Survival Analysis in Longitudinal Studies for Recurrent Events: Applications and Challenges.” *Clinical Epidemiology and Global Health* 7 (2): 253–60. https://doi.org/10.1016/j.cegh.2019.01.013.
14. Therneau, Terry M., and Patricia M. Grambsch. 2000. *Modeling Survival Data: Extending the Cox Model*. *Statistics in the Health Sciences*. https://doi.org/10.1007/978-1-4757-3294-8.
15. Villeneuve, Andréanne, Valérie Arsenault, Jacques Lacroix, and Marisa Tucci. 2020. “Neonatal Red Blood Cell Transfusion.” *Vox Sanguinis* 116 (4): 366–78. https://doi.org/10.1111/vox.13036.
16. World Health Organization: WHO. 2024. “Newborn Mortality.” March 14, 2024. https://www.who.int/news-room/fact-sheets/detail/newborn-mortality.
17. Xi, Chuhao, Min Luo, Tian Wang, Yingxiang Wang, Songbai Wang, Lan Guo, and Ciyong Lu. 2020. “Association Between Maternal Lifestyle Factors and Low Birth Weight in Preterm and Term Births: A Case-control Study.” *Reproductive Health* 17 (1). https://doi.org/10.1186/s12978-020-00932-9.