College of Natural and Computational Sciences

Department of Statistics

Competing Risk Analysis of the Health Status of Neonates with Respiratory Distress Syndrome

By: Zenaw Ayele

A Thesis Submitted to the Department of Statistics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Biostatistics)

Advisor: Mekonnen Tadesse (Assoc. Prof)

Addis Ababa University
Addis Ababa, Ethiopia
June 2020

Acknowledgment

Next to God for his unlimited blessing in my entire life, my deepest gratitude goes to my advisor Mekonnen Tadesse (Assoc. Prof) for his valuable contribution to accomplish this thesis. His expert guidance, unlimited assistance, and constructive comments have been vital in all phases of the thesis work.

I gratefully thank Mr. Zelalem Tazu for his support, encouragement, and unlimited advice from title selection to the final report of this study. My sincere thanks go to all staff members of the Department of Statistics at Addis Ababa University especially, my lecturers Dr. Merga Belina, Dr. Dejen Tesfaw, Prof. M. K. Sharma, and Dr. Bedilu Alamirie. I would also like to express my special appreciation and thanks to Professor Eshetu Wencheko for providing a very important tutorial on the basics of thesis writing which helped to perform this thesis.

Above all, I wish sincere thanks to the Ethiopian Statistical Association and the current executive members of the association for allowing me to pursue my graduate studies and their unlimited support. Finally, I am so glad to thank my family and my friend Mr. Dagnachew Fenta for their encouragement and support throughout my life.

ABSTRACT

Respiratory distress syndrome (RDS) is not only the most common respiratory disorder in premature infants but also the main cause of neonatal mortality. Data from a study of illness in preterm project collected from neonatal intensive care units (NICUs) of five selected hospitals in Ethiopia during July 1, 2016, to May 31, 2018, were used to examine and identify potential prognostic factors of the health status of preterm infants with respiratory distress syndrome by considering the competing risk framework. Preterm infants admitted to NICU of the selected hospitals were followed for 28 days and only neonates with complete cases and RDS problems were included in the analysis. The Fine-Gray or sub-distribution hazard model was used to identify significant prognostic factors. Three outcome variables (death due to RDS, death due to other causes and discharged alive) were considered and the Fine-Gray model was fitted for each outcome variable separately. Anemia, multiple pregnancies, birthweight, and gestational age were the potential prognostic factors significantly associated with the death of neonates due to Respiratory distress syndrome problem. Pneumonia, meningitis, anemia, and gestational age of neonates were the significant prognostic factors for the death of neonates due to other causes. Moreover, pneumonia, birthweight, and gestational age were identified as the prognostic factors associated with being discharged alive of the neonates. Offering intensive and adequate treatments for most critically exposed neonates with lowest birth-weight (less than 1000g) and gestational age (less than 28 weeks) may be useful to decrease the burden of neonatal mortality and increase the incidence of being discharged alive.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

CI	Confidence Interval
CIF	Cumulative Incidence Function
CR	Competing Risk
CSH	Cause-Specific Hazard
DOC	Death Due to Other Causes
EI	Event of Interest
GMH	Gandhi Memorial Hospital
GUH	Gonder University Hospital
JUH	Jimma University Hospital
KM	Kaplan Meier
NICU	Neonatal Intensive Care Unit
PH	Proportional Hazard
RDS	Respiratory Distress Syndrome
SD	Standard Deviation
SDGs	Sustainable Development Goals
SDH	Sub-distribution Hazard
SIP	Study of Illness in Preterm
SPH	Saint Paul's Hospital
TAH	Tikur Anbessa Hospital

1. INTRODUCTION

1.1. Background of the Study

Respiratory distress syndrome (RDS) is one of the neonatal health complications which require a lot of tasks to be targeted for the development of treatments and technology for neonatal intensive care (Kamath et al., 2011). It frequently happened in premature infants and is the most common cause of admission to the neonatal intensive care units (NICUs), with the clinical symptoms like tachypnea, poor feeding, nasal flaring, grunting, cyanosis, intercostal retraction and reduction of respiratory sounds in pulmonary auscultation (Wang et al., 2015). Moreover, RDS is considered as the most common respiratory disorder in premature infants and mainly prematurity increases the vulnerability of neonates to develop RDS (Hoxha et al., 2018). Globally, RDS is the main cause of neonatal mortality. But, information about RDS-specific mortality rates and technologies for its treatment are rare in low-income countries leading towards inconsistent health care platforms (Kamath et al., 2011).

The neonatal mortality rate continues to be an urgent global problem about 2.5 million infants died in the first 28th days of life, the risk of dying was highest in the first month of life (UNICEF, 2018). Globally, neonatal death takes about 47% of all deaths of under-five children. Moreover, more than one- third of preterm deaths occurred due to preterm birth complications, which is the leading cause of death and respiratory distress syndrome is one of the primary causes of death of neonates (Muhe et al., 2019). There are different policies, strategies, and programs that work on prevention and care of preterm birth and its outcome including RDS. Among these are Sustainable Development Goals (SDGs) and Every woman and Every Child initiative. However, the etiologies and risk factors associated with RDS have not been well-cited in low-income countries and particularly in sub-Saharan Africa (Aluvaala et al., 2019).

The availability of adequate treatments and having a good understanding of RDS and its consequences could prolong the survival time of neonates. In this study, we have focused on a preterm infant with RDS problem which is the most common cause of morbidity and mortality of neonates admitted to the neonatal intensive care unit (NICU). In our case neonates admitted to the NICU due to the problem of RDS were followed for 28 days and whenever a death occurred the primary causes of death were assessed. Though they were admitted to NICU due to

respiratory distress syndrome problem, some neonates died due to other causes or discharged alive from the NICU in the follow-up days. Thus, the occurrence of neonatal death due to other causes or of those discharged alive from the NICU precludes the observation of death due to RDS. As a result, the existence of two or more than two outcome events in which the occurrence of one event prevents the occurrence of the event of interest leads to a special approach called competing risk analysis. Customarily, prognostic factors of neonatal data were analyzed using the Cox proportional hazards model by considering the outcome events as either event of interest or censored observations though, during follow-up, neonates in the neonatal intensive care unit often died from other causes rather than the event of interest, such as cardiovascular disease, maternal issues, and other unexpected events. In the standard survival models, such neonates are treated as censored observations which may lead to biased estimates.

A competing risk is an event that either hinders the observation of the event of interest or modifies the chance of the occurrence of this event (Noordzij et al., 2013). Equivalently, competing risk refers to a situation where an individual is exposed to two or more causes of failure, and its eventual failure can be attributed exactly to only one (Pintilie., 2006). In competing risk analysis, there is a distinct cause-specific hazard function for each of the distinct types of events and a distinct sub-distribution hazard function for each of the distinct types of events (Austin et al., 2016). In this study, the events that may be observed, called competing risks, include the death of neonates due to respiratory distress syndrome (RDS) which is our event of interest, death due to other causes (DOC), and those discharged alive from hospital. Thus, we aim to examine and identify accurate prognostic factors associated with the health status of neonates by considering competing risk framework.

1.2. Statement of the problem

The existence of adequate measurement for neonatal mortality in health centers is essential to improve the quality of health services and it is important to assess the achievement of sustainable development goal three (SDG 3) in developing countries (Aluvaala et al., 2019).

Studies conducted by Liu, Oza, Hogan, et al. (2016), Muhe et al. (2019), Yismaw et al. (2019), and others have identified respiratory distress syndrome as a primary cause of death of neonates in neonatal intensive care units (NICUs). However, most previous works done on the issue had applied the standard survival analysis methods without considering the occurrence of other

events which could potentially alter the chance of the occurrence of the event of interest. But ignoring the competing event(s) and using ordinary survival analysis methods such as the Kaplan–Meier method and standard Cox proportional hazards regression, may be inappropriate in the presence of competing risks, and alternative methods specifically designed for analyzing competing risks data should then be applied (Noordzij et al., 2013). Moreover, thought neonates were admitted to NICU due to RDS, some neonates died due to other causes or discharged alive which could either prevents observation or probability of occurrence of the event of interest, death due to RDS. Thus, using standard survival analysis methods would lead to censoring neonates who died due to other causes and discharged alive which may be resulting in a statistical error (maybe the wrong magnitude of parameter estimates) and consequently leading to incorrect conclusions. These limitations could be handled using competing risk models/analysis methods Therefore, the study aimed to obtain potential prognostic factors through appropriate modeling approaches by accounting for the potential competing risks. Moreover, to the best of our knowledge, there were no previous works on competing risk analysis of neonates with RDS in Ethiopia.

1.3. The objective of the Study

This study aims to examine and identify the prognostic factors associated with the health status of neonates with RDS.

1.4. Significance of the Study

This study presented important methods appropriate for examining the prognostic factors in the presence of competing events where estimates were adjusted and possible biases controlled. Consequently, the results of this study are expected to provide information about prognostic factors of death of neonates (with RDS) admitted to NICU by considering competing events that may in turn be essential for planning effective programs and evaluating the existing national health policy. Besides, the results of this study may be used as a basis for future studies in the area.

1.5. Limitations of the Study

One of the limitations of our study is not being able to make sufficient comparisons of the findings of this study with findings of other researchers due to lack of previous work on the health status of neonates (with RDS) which accounted competing risk events. Besides, the data

used in this study were obtained from the study of illness in preterm project and data from only five selected hospitals were considered making the result of the study not representative of the national level situation.

1.6. Organization of the Thesis

The remaining part of this thesis is structured as follows. In chapter 2, we have presented a review of related literature about neonatal mortality due to respiratory distress syndrome, models suggested for data on neonates admitted to NICU, risk factors for neonates with RDS, and literature on competing risk models. In chapter 3, we have emphasized the data source, competing risk model for survival data, methods of parameter estimation, and model assumption checking. In chapter 4, we presented the distribution of neonatal characteristics, descriptive, and analytical results. In chapter 5, we have discussed the results obtained in chapter 4. In the last part of this thesis (chapter 6), we have made conclusions and recommendations.

2. LITERATURE REVIEW

A cross-sectional study done on the prevalence of RDS in neonates in Saudi Arabia by Qaril et al. (2018) indicated that RDS was one of the major problems among newborns and a major cause for increased disease and death among preterm infants. Moreover, they have reported that hypertensive disorder, diabetic mothers, Cesarean delivery, and being male were the important risk factors for RDS. Also, a retrospective cohort study conducted by Luo et al. (2019) on the prevalence and predictors of infant mortality due to RDS in Beijing and Jiangxi Provinces of PR China, applied multivariate logistic regression analysis methods. The results of their study indicated that infants with extremely low birth-weight (< 1000g), gestational age (< 37 weeks), and cesarean delivery were the independent risk factors of infant death due to RDS.

A hospital-based observational study in India conducted by Patil et al. (2018) on the risk factors associated with the development of severe respiratory distress in the newborn had used the t-test and Chi-square test. Their study indicated that high maternal age, mode of delivery, parity, and birth weight were risk factors associated with severe respiratory distress in newborn infants. In particular, their study showed that the risk of severe respiratory distress was higher for low birth weight. Similarly, a multi-center observational study was done by Wang et al. (2019) on risk factors of RDS among Chinese infants of 34-42 weeks gestational age, and logistic regression analysis was implemented to compare the risk factors for respiratory distress syndrome at different gestational ages. Their finding revealed that an Apgar score < 7 at 5 min post-birth, gender, and diabetic mothers are independent risk factors of RDS for late preterm and term of the study area. For infants with gestational age 34-38 weeks, the incidence of RDS decreased with advancing gestational age. However, cesarean delivery had a negative effect or reduced the incidence of RDS for late preterm infants.

A population-based study in Sweden done by Altman et al. (2013) on the risk factors for acute respiratory morbidity in moderately preterm infants had used multivariate logistic regression analysis. Their finding showed that multiparty, cesarean section, sex, Apgar score <7 at 5 minutes, and lower gestational age were risk factors of respiratory distress syndrome for premature infants. Specifically, the preterm rupture of membranes, antenatal corticosteroid treatment, and being small for gestational age reduced the risk of RDS. Besides, a study conducted in Italy by Condò et al. (2016) on neonatal respiratory distress syndrome to

investigate whether the risk factors are the same in preterm and term infants carried out multivariate analysis by categorizing newborns as infants in early and moderate preterm, late preterm and term infants. The results of their study showed that low birth weight was a risk factor for term infants, mode of delivery via cesarean section, and being male are risk factors for the incidence of RDS in all gestational age. But, maternal age and multiple births were associated with the incidence of RDS in any group.

A retrospective investigation conducted in Poland by Hoxha et al. (2018) on the risk of RDS in singleton pregnancies with preterm premature rupture of membranes between 24+0 and 36+6 weeks had applied logistic regression model. The main RDS risk factors for premature neonates were gender, abnormal placental circulation, and fetal distress.

Tochie et al. (2016) investigated the prevalence, predictors, etiologies, and outcomes related to RDS in a reference neonatal unit in Cameroon. They applied the logistic regression model and the results of their study revealed that RDS is a significant cause of high neonatal morbidity and mortality in developing countries. Furthermore, hospital-based studies in Africa on the burden and predictors of neonatal mortality reported that preterm accounts about 15.7 to 29.6%. Studies in Ethiopia on causes, survival, predictors, and implications of preterm neonatal mortality reported the rate to be ranging from 18% to more than 40%. RDS was identified as one of the neonatal clinical problem-related risk factors for preterm neonatal survival (Yismaw et al., 2019). A prospective, cross-sectional observational study of neonates in five selected hospitals in Ethiopia also identified respiratory distress syndrome as a primary cause of death of neonates (Muhe et al., 2019).

Alfarwati et al. (2019) studied the incidence, risk factors, and outcome of Respiratory Distress Syndrome in term infants at the academic Centre, Jeddah, Saudi Arabia using t-test and Chi-square test. Their finding indicated that low birth weight and low Apgar scores were the main risk factors of RDS for full-term neonates. Also, Asmare et al. (2019) studied the incidence of respiratory distress and its predictors among neonates admitted to the neonatal intensive care unit of Black Lion Specialized Hospital, Addis Ababa, Ethiopia using standard survival analysis methods. The results of their study revealed that neonates delivered at home, delivered through

cesarean section, preterm neonates, whose APGAR score < 7, and born from diabetic mothers were more likely to develop respiratory distress syndrome.

A multivariable logistic regression analysis of the treatment patterns and clinical outcomes in neonates diagnosed with RDS in a low-income country was carried out by Hubbard et al. (2018). They examined the associations between potential risk factors and the primary outcome of death before discharge. The result of their study revealed that place of birth, mode of delivery, gestational age, and birth weight were significant risk factors for death of neonates.

Investigation on causes and outcomes of respiratory distress in neonates hospitalized in the neonatal intensive care unit of Be'sat Hospital in Hamadan, Iran used logistic regression and the results of their study showed that RDS is the most common cause of respiratory distress in the hospitalized neonates (Sabzehei et al., 2017).

A study conducted by Gargari et al. (2017) on survival and risk factors of extremely preterm babies (< 28 weeks) in the three Iranian Hospitals used multiple logistic regression analysis and reported that younger maternal age, lower neonatal birth weight, and gestational age can increase the risk of neonatal mortality. Moreover, their finding indicated that receiving corticosteroid and being female could increase neonates' survival time and it had a positive significant relationship with the survival rate of infants (P<0.05).

A retrospective cohort study by Orsido et al. (2019) used the Kaplan-Meier method to show the pattern of death in 28 days and the Cox-proportional model to identify the predictors of neonatal mortality in neonatal intensive care unit at a referral hospital in Southern Ethiopia. They reported that the neonatal mortality incidence was 27 per 1000 neonates-days and parity, attend antenatal care follow up, mode of delivery, initiated breastfeeding, neonates resuscitated, hyaline membrane disease and perinatal asphyxia were significant predictors of neonatal mortality.

Wesenu et al. (2017) used non-parametric, semi-parametric, and parametric survival models to identify the determinants of time-to-death of premature infants admitted to the Neonatal Intensive Care Unit in Jimma University Specialized Hospital. The estimated survival times and investigated the association between survival times and different potential variables. Their finding showed that low gestational age, respiratory distress syndrome, and initial temperature

were significant factors for neonatal mortality. Standard survival analysis methods were applied by Yismaw et al. (2019) in their study on survival and predictors of death among preterm neonates admitted at the University of Gondar Comprehensive specialized hospital neonatal intensive care unit, Northwest Ethiopia. The results of their study revealed that sex, birth weight, gestational age (GA) at birth and neonatal congenital malformations, neonatal clinical problems like respiratory distress syndrome (RDS), perinatal asphyxia (PNA, jaundice, hypoglycemia, hypothermia and neonatal sepsis and timely initiation of breastfeeding upon birth are significant factors of preterm neonatal death.

A study on time to death or discharge in neonatal care by Hinchliffe et al. (2013) applied competing risk analysis methods and their result indicated that competing risks were a suitable statistical method for modeling length of stay in the presence of significant rates of in-hospital mortality. Besides, the discharge of babies alive generally occurred over a longer time for babies of lower gestational age and smaller birthweight than for bigger babies. In their report, they have shown that the competing risks are a suitable statistical method for modeling length of stay in the presence of significant rates of in-hospital mortality.

A study by Aluvaala et al. (2019) studied on competing risk survival analysis of time to in-hospital death or discharge in a large urban neonatal unit in Kenya have applied competing risk modeling approach and their finding indicated that the probability of in-hospital death of neonates was higher than that for neonates discharged alive for birth-weights less than 1.5 kg with the transition to a higher probability of being discharged alive observed after the first week in birth-weights 1.5 to <2 kg.

According to the results of a study done by Van Walraven et al. (2015), almost half of Kaplan-Meier estimates obtained from standard survival analysis of studies published in prominent medical journals may overestimate the risk of the event of interest due to the occurrence of competing events. Moreover, they have suggested that researchers should avoid using ordinary survival analysis methods to estimate risks when there are competing events and that the cumulative incidence function-based estimates should be used in such cases.

A study conducted by Resche-Rigon et al. (2005) that focused on modeling the mortality of patients in the intensive center unit suggested that data from intensive units be analyzed based on a competing risk.

3. DATA AND METHODS OF ANALYSIS

3.1.Data Sources

The data considered in this study were obtained from Study of Illness in Preterm (SIP) project that compiled neonatal data from five selected hospitals (Gondar University Hospital, Jimma University Hospital, Gandhi Memorial Hospital, Saint Paul's Hospital Millennium Medical College, and Tikur Anbessa Hospital) in Ethiopia which have Neonatal Intensive Care Unit (NICU). The primary objective of the SIP project was to determine the most common causes of illness and mortality in preterm infants admitted to the selected hospitals in Ethiopia. The study was conducted in Ethiopia and funded by external grants from the Bill & Melinda Gates Foundation. The project had its data management system (DMS) which was developed by Addis Ababa University and used at each of the participating study sites. Electronic data were transferred weekly from each data management computer to the central data center at Addis Ababa University, creating a complete data repository. Records from the study site were transferred to the central server found in Addis Ababa University at Tikur Anbesa Hospital and then merged with the existing database to add new records and update edited records and the data for this study was obtained the from this hospital. The neonates were recruited to the study over nearly 2 years (from July 1, 2016, to May 31, 2018) admitted neonates were followed for 28 days. Also, the study subjects were preterm infants born before 37 completed weeks of gestation and admitted to neonatal intensive care units (NICUs) of one of the selected hospitals due to the problems of RDS. Besides, consent was made with parents or caregivers for post-mortem examinations and whenever a death occurred, the primary cause and date of death of the neonate was recorded.

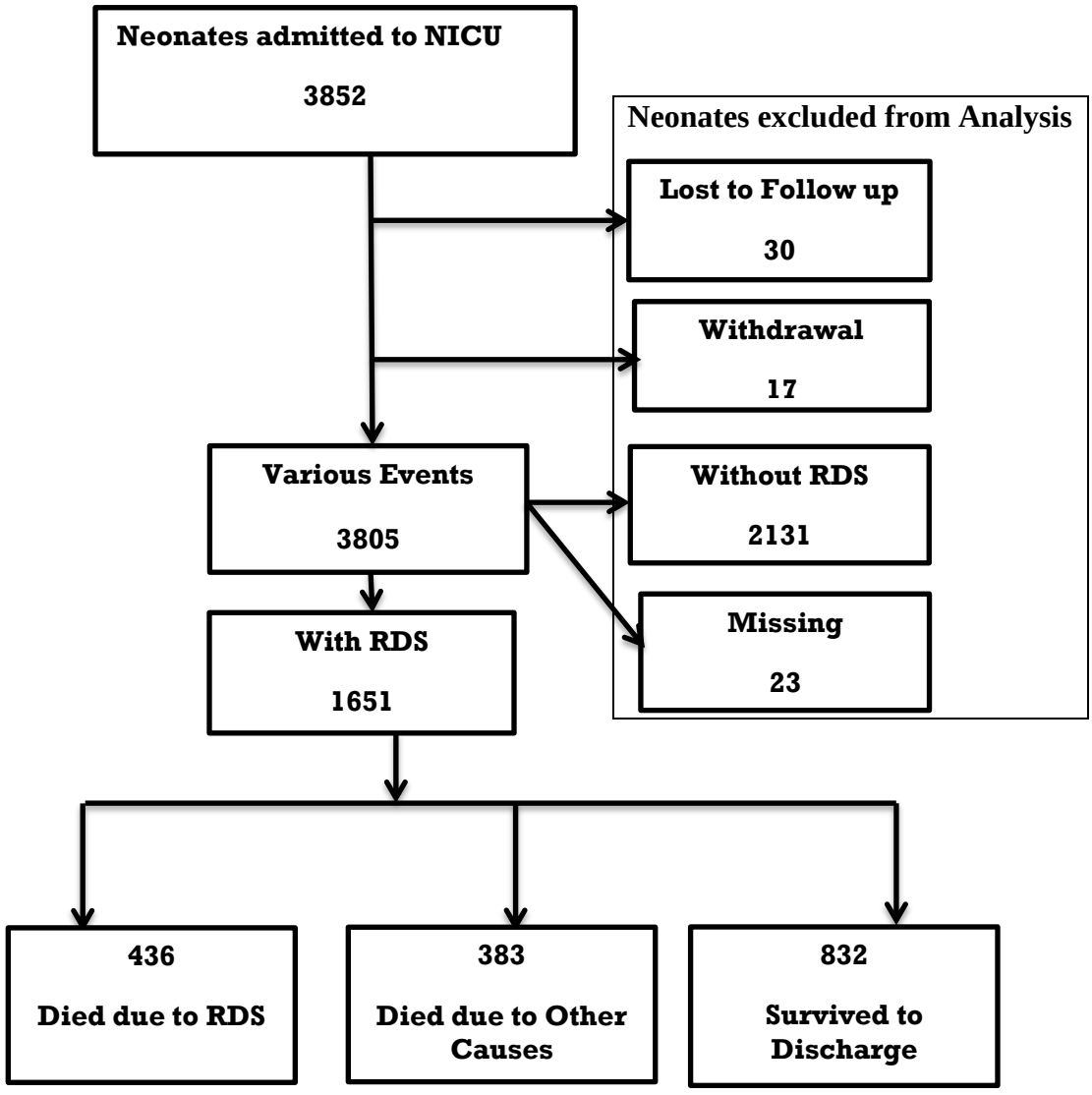
3.2.Variables used in the Study

Based on the literature review, the covariates considered in the study are the gender of preterm infant (male, female), gestational age in weeks (<28, 28-31, 32-34,>35), whether mothers had multiple pregnancies (yes, no), birth-weight of neonates at the time of birth in gram (<1000, 1000-1500, 1500-2000,>2000), mothers' age at birth in years (<20,20-34,>=35), having pneumonia (yes, no), having anemia (yes, no), the existence of feeding problem with partial to total sucking refusal or difficulty sucking (yes, no), mother has received antenatal care (yes, no), mother diabetic (yes, no), C-section during delivery (yes, no), hypertensive disorder during

pregnancy (yes, no). We have considered three outcome variables: death due to RDS, death due to other causes (i.e. neonates admitted to NICU due to RDS died due to other causes such as feeding problem and maternal diabetic nature) and discharged alive from the neonatal intensive care unit. While the death of neonates due to RDS was our event of interest and death due to other causes and neonates discharged alive from the unit are competing risks that compete with the event of interest. The outcome variable (status) takes numeric values with 1= died from RDS, 2= died from other causes, and 3=survived to discharge from NICU. Since neonates in the study experienced one of the possible outcomes so there are no censored observations in the current study. All neonates with the health problem of respiratory distress syndrome during the study period were included in the study (inborn in the study hospital and referral from other hospitals within 2days of birth). A neonate with incomplete data of its medical records was excluded from the analysis.

In this study, a total of 819 preterm infants died until the end of the follow up of which 436 neonates died due to RDS accounting for 53.2% of the deaths and 26.2% of the total preterm neonates admitted to the neonatal care unit due to the problem of RDS. Moreover, 383 patients died due to other causes accounting for 46.8% and 832 neonates (50.4% of the total admitted neonates) were discharged alive from the hospitals. The detailed screening process of the neonates that were considered in this study is shown in Figure 3.1.

Figure 3.1: Screening of Neonates Considered in the Study



3.2.1. Coding and Description of Variables

Table 3.1. Coding of Variables considered in the study

No	Variables	Categories
1	Gender	(0) Female (1) Male
2	Gestational Age (Weeks)	(0) <28 (1) 28-31 (2) 32-34 (3) 35 and above
3	Multiple pregnancy	(0) No (1) Yes
4	Birth weight of neonates (gram)	(0) <1000 (1) 1000-1500 (2) 1500-2000 (3) 2000 and above
5	Feeding Problem	(0) No (1) Yes
6	Pneumonia	(0) No (1) Yes
7	Maternal Age	(0) <20 (1) 20-34 (2) >=35
8	Antenatal Care Received	(0) No (1) Yes
9	Cardiac Disease	(0) No (1) Yes
10	Diabetic mother	(0) No (1) Yes
11	Hypertensive Disorder	(0) No (1) Yes
12	C-section	(0) No (1) Yes
13	Anemia	(0) No (1) Yes
Outcome variables		
No	**Variable**	**Description**
1	Follow up time (days)	Length of stay of the neonates in the hospital
2	RDS	Neonates died due to Respiratory Distress Syndrome
3	Other cause of disease	Neonates died due to other causes
4	Discharged alive from Hospital	Survived, discharge from NICU

3.3.Competing Risk Analysis of Survival Data

A competing risk is an event in which the occurrences of an event prevents the occurrence or fundamentally alter the probability of the event of interest (Austin et al., 2016). It encountered in studies where the subjects under study are at risk of more than one mutually exclusive event or failure causes. For example, in a follow-up study of patients with cardiovascular disease, a patient may die due to cancer, and cancer is said to be a competing risk since the patient died due to cancer and not due to cardiovascular disease. Besides, ignoring the potential effects of

competing events can easily lead to incomplete reporting, incorrect risk estimates, and potentially wrong interpretation.

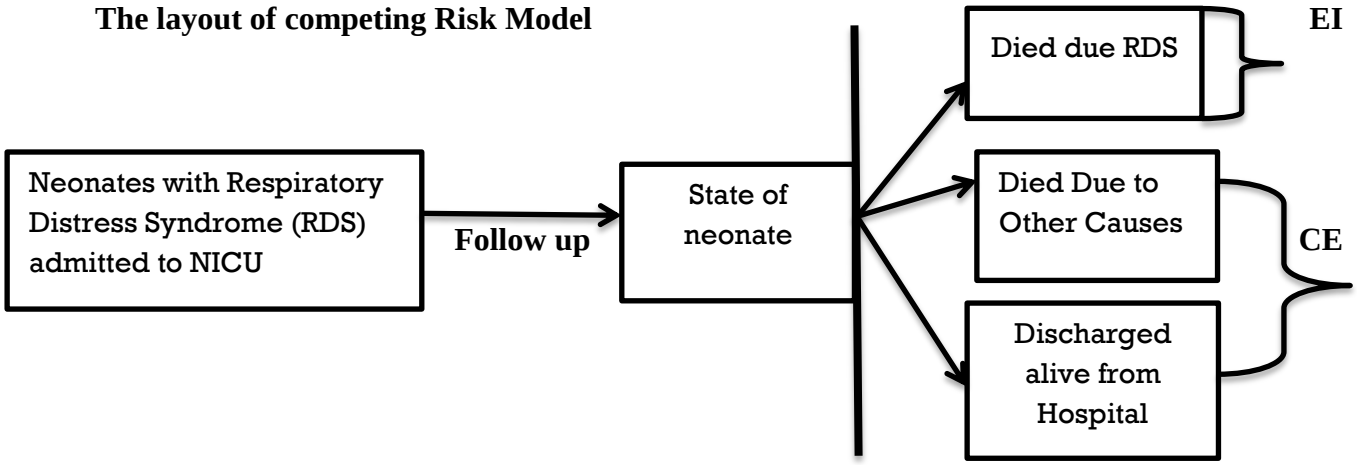
In fitting models in the presence of competing risks, one can choose from two different modeling approaches: cause-specific hazard or Sub distribution hazard model also known as the Fine-Gray Model. The cause-specific hazard model is used to estimate the effect of the covariates on the rate of occurrence of the outcome in those subjects who are currently event free. But the Fine-Gray Model allows us to estimate the effect of covariates on the absolute risk of the outcome over time.

Moreover, standard survival methods are not appropriate to analyze the survival data in a competing risk setup where an individual in the risk set is exposed to more than one cause for failure. Among the most popular methods for analyzing the competing risk data is the Cox proportional hazard model (Cox, 1972) to examine the effect of covariates on the cause-specific hazard function. The major limitation of using this model for competing risk data is that during the estimation of regression parameters under a specific cause, it considers the individuals failing from causes other than the cause of interest as censored observations. To overcome the limitation, Fine and Gray (1999) designed a new approach based on cumulative incidence function (CIF) which describes the probability of occurrence of an event before a specific time. In contrast to the Cox model, using CIF based model does not ignore the other competing risks when a specific cause is of interest. Besides, the Fine-Gray model is based on a proportional hazards model for the sub-distribution of a competing risk, where covariates under the study directly affect the cumulative incidence function.

The CIF for the k^{th} failure or cause is given by $Pr(T \leq t, D = k)$, where D is the type of event that happened from the possible event at survival time t. In standard survival analysis, we know that the survival function is given by $S(t) = P(T \geq t)$, and the incidence of the event throughout follow-up is given as $F(t) = 1-S(t) = Pr(T \leq t)$. From this, we can see that the cumulative incidence function (CIF), as distinct from $1 – S(t)$, allows for estimation of the incidence of the occurrence of an event while taking competing risk into account. This allows one to estimate incidence in a population where all competing events must be accounted for in clinical decision making.

A main point is that, in the competing risks setting, only one event type can occur, such that the occurrence of that event precludes the subsequent occurrence of other event types. The function CIF_k (t) denotes the probability of experiencing the k^{th} event before time t and before the occurrence of a different type of event. Unlike the survival function in the absence of competing risks, CIF_k (t) will not necessarily approach unity as time becomes large because of the occurrence of competing events that preclude the occurrence of events of type k. Thus, the CIF or sub-distribution function could be greater than one and hence it is not a proper distribution (Pintilie, 2006).

The layout of competing Risk Model



Where EI refers to the event of interest and CE refers to the competing events

3.4.Basic Hazard Functions in Competing Risk Analysis

There are two common types of hazard functions in competing risk analysis. These are the cause-specific hazard and sub-distribution hazard (cumulative incidence functions)

The cause-specific hazard function for the k^{th} cause, (Austin et al., 2016) is defined as

$$\lambda_k{}^{cs}(t)=\lim_{\Delta t\to 0}\frac{\text{Prob}(t\le T<t+\Delta t, D=k|T\ge t)}{\Delta t} \quad k=1,\ldots,D \qquad (1)$$

and represents the rate of occurrence of the k^{th} failure. In our case k=1, 2, 3

It denotes the instantaneous rate of occurrence of the k^{th} event in subjects who are currently event-free (i.e. in subjects who have not yet experienced any of the different types of events).

The sub-distribution function for the k^{th} event, (Austin et al., 2016) is defined by

$$\lambda_k{}^{sd}(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}(t \leq T < t+\Delta t, D=k | T > t \cup (T < t \cap D \neq k))}{\Delta t} \quad (2)$$

$$F_k(t) = P(T \leq t, D = k) \qquad k=1,\ldots,D$$

and corresponds to the probability of a subject failing from cause k in the presence of all the competing risks. The CIF is used to model the risk of experiencing a specific event in subjects who have not yet experienced this event. It denotes the instantaneous risk of failure from the k^{th} event in subjects who have not yet experienced an event of type k. The basic difference between the two hazards is related to the risk sets. The risk set is the set of individuals /subjects under investigation and vulnerable to the event. In CIF, the risk set includes those who are currently alive as well as those who have previously experienced a competing event while the risk set of the cause-specific hazard function only considers those who are currently event free.

For univariate analysis of each prognostic factor, cumulative incidence function was used and the value of Gray's test was obtained to determine the extent of a significant association between the outcome variable and the factors. Gray's test is a K-sample test that was introduced by Gray (1988). It compares the weighted averages of the sub-distribution hazards across groups for the event of interest. The null hypothesis is that there is no difference across the given groups.

3.5.Regression Models of Competing Risks

In survival analysis with competing risks, two regression modeling approaches depend on the above-mentioned hazard functions: the cause-specific hazards model and the Sub-distribution hazard model also called the Fine-Gray model (Porta et al., 2007). The cause-specific hazard model is used to estimate the effect of the covariates on the rate of occurrence of the outcome in those subjects who are currently event free. In contrast, the Fine-Gray Model allows us to estimate the effect of covariates on the absolute risk of the outcome over time Fine-Gray (1999). However, there are also some advanced modeling approaches used in competing risk modeling indicated in the literature such as mixture models, vertical modeling and the analysis of time-to-event data based on pseudo-observations (Haller et al., 2013).

3.5.1. Cause-specific Hazard Model

In modeling the cause-specific hazards, each hazard is analyzed separately by treating individuals failing from other causes as censored observations. This approach is appropriate

when interest is to determine which factors affect the rate at which events happen. The usual regression analysis of competing risks establishes a Cox proportional hazards (PH) model (Prentice et al., 1978) for each cause-specific hazard and defined as:

$$\lambda_k(t|X) = \lambda_{0k}\exp(X\beta_k) \qquad k=1,\ldots, D \qquad (3)$$

Where X is a px1 vector of covariates, β_k is a px1 vector of unknown parameters to be estimated for each outcome variable, k is the cause of failure at time t. And λ_{0k} are the baseline cause-specific hazard and the PH assumption, in this case, is a strong one that must be carefully checked for each cause.

3.5.2. Fine-Gray Model

The Cox PH model has been commonly applied to analyze competing risks data. Since this method has several limitations in analyzing survival data in the presence of competing risks, Fine and Gray (1999) developed a CIF based proportional hazard model to analyze competing risk data. In the competing risks set up, for each cause, the occurrence of an event of interest, a hazard function in the presence of covariates are considered (Mohammad et al., 2017). It is useful when interest is to know which covariates affect the probability of an event occurring over time. Moreover, the sub-distribution hazard ratio obtained from the Fine-Gray model indicates the relative effect of covariates on the sub-distribution hazard function or cumulative incidence function (Austin et al., 2017). The sub-distribution hazard model is also used to determine factors associated with the incidence of a given cause. This method of analysis does not treat individuals failing from other causes as censored observations.

In the presence of competing risks, the choice of competing risk model depends on the objective of the study. Scholars suggested that use the sub-distribution hazard model when the focus is on estimating incidence or predicting prognosis in the presence of competing risks and use the cause-specific hazard model when the focus is on addressing etiologic questions (Austin et al., 2016). In this thesis, we have applied the Fine-Gray modeling approaches to identify the potential prognostic factors of the health status of neonates that have respiratory distress syndrome problems.

Using the relationship between the survival, hazard and cumulative incidence function (Leoce, 2016)

$$\lambda(t)=\frac{f(t)}{S(t)}=\frac{f(t)}{1-F(t)} \qquad (4)$$

The sub-distribution hazard (hazard of the cumulative incidence) for each cause supposed as the hazard for an individual who either fails from cause k or does not, can be written as:

$$\lambda_k{}^*(t;X) = \frac{f_k(t)}{1-F_k(t)} \qquad (5)$$

Under the proportional hazard, Fine-Gray Model can be specified (Leoce, 2016) as:

$$\lambda_k{}^*(t; X) = \lambda_{0k}{}^*(t) \exp(X\beta_k) \qquad (6)$$

Where $\lambda_{0k}{}^*(t)$ is the baseline sub-distribution hazard for the cause of k and exp (β_k) is the relative risk probability of k^{th} cause associated with the given covariates.

3.6.Variable Selection Methods in Competing Risks

When fitting regression models, the decision about which of the covariates to be included in the model is very important. There are several variable selection methods used in different modeling approaches. However, in competitive risk analysis variable selection is not straightforward. The stepwise methods are the most common methods for selecting variables and it includes forward selection, backward selection, or a combination of both. The Bayesian information criterion (BIC), Akaike information criterion (AIC), and BICcr (modification of BIC to competing risks models are the available methods for variable selection in competing risk analysis). The AIC is -2logL(β) +2p and the BIC is -2logL(β) + p*log*n, where p is the number of parameters in the model, n is the number of observations and *log*L(β) is the maximum value of the (weighted) log partial likelihood for the Fine and Gray model (Kuk et al. ,2013).

In the Fine and Gray model, only subjects who experience the primary event contribute information to the partial likelihood. According to the work of Kuk et al. (2013), they proposed a new variable selection criterion called BICcr by changing the penalty to be the total number of primary events and denoted as n*.

The equation for BICcr is defined as:

$$\text{BICcr} = -2 \log L(\beta) + p\log(n^*). \qquad (7)$$

However, in our analysis, we have used the full range of potential prognostic factors through careful study of the literature concerning the outcome variable. For each outcome variables (death due to RDS, death due to other causes and discharged alive, the Fine-Gray model was fitted separately.

3.7.Methods of Parameter Estimation

Estimation of parameters in the Fine-Gray model uses the partial likelihood approach similar to the standard Cox model since a proportional hazard assumption is imposed on the sub-distribution hazards (Pintilie, 2006). However, in this model, the parameters are estimated by incorporating weights in the partial likelihood.

The partial likelihood for the Fine-Gray model is given (Kuk et al., 2013) as:

$$L(\beta)= \prod_{k=1}^{r} \frac{\exp(X\beta_k)}{\sum_{i\in R^*(t_k)} w_{ki} \exp(X\beta_i)} \qquad (8)$$

The product is taken over all r time points, ($t_1 < t_2 < \ldots < t_r$), where r is the total number of the event of interest i.e. ($\sum_{i=1}^{n} \text{Indicator}\{\in_i = 1\}$). The modified risk set,$R^*(t_k)$ is a set of subjects that are still at risk for the event of interest at time t (i.e., those who did not experience the primary event and are not censored by time t) Fine-Gray (1999). Thus, subjects that have experienced other types of events remain in the risk set all the time. Besides, the weight is defined as

$$w_{ki} = \frac{\hat{G}(t_k)}{\hat{G}(\min(t_k,t_i))} \qquad (9)$$

where $t_i = \min(T_i,C_i)$ for i such that $\in_i \neq 1$ and t_k is the time of the k^{th} event. $\hat{G}$ is the KM estimate of the survivor function of the censoring distribution ($G(t) = P(C \geq t)$). The weight is one for the subjects who did not experience any type of event by time t_k and less than 1 for those who had a competing event before t_k. As a result, individuals who experience a competing event at time t_i do not participate fully in the partial likelihood; the further the time point (t_k) is from the time of the competing event (t_i) the smaller the weight. When there is only one event of interest, the weights are all equal to 1, and the risk set contains only those at risk at the specified time point (Kuk et al., 2013). Once the likelihood is formulated, the goal was to choose the values of parameters that maximize the likelihood.

3.8.Model Diagnosis

The main assumption when modeling survival data is the proportionality of hazards. When the Fine-Gray model is used, the hazards of the CIF must be proportional whereas, in the Cox proportional hazard model, it is the cause-specific hazards that need to be proportional (Pintilie, 2006). The proportionality assumption is the most common in competing risk regression model,

which considers the sub-distribution with covariates X is a constant shift on the complementary log-log scale from a baseline sub-distribution function. If the curves do not cross with each other then we say that the model does not violate the assumption of proportionality (Zhang, 2017).

3.8.1. Proportionality of the cause-specific hazards

A visual inspection of the plot of log (−log (S)) vs log (time) can give a hint as to whether the cause-specific hazards can be assumed to be proportional. In this case, S is the Kaplan–Meier estimate when the only event considered in the event of interest; both the observations without an event, as well as the competing risks, are censored. Such a plot can be drawn for each of the levels of a covariate (Pintilie, 2006).

3.8.2. Proportionality of the hazards of the CIF

To investigate the proportionality assumption for the competing risks regression, log (−log (1−F)) can be plotted against log (time) where F is the CIF for the event of interest (Pintilie, 2006). In this work, we have applied the Fine-Gray model and we have checked the assumption of proportionality assumption of the CIF.

4. RESULTS

4.1.Descriptive Results

A total of 1651 eligible neonates were enrolled in the centers due to the problem of RDS. Thus, the data considered for this study was neonatal data admitted to the NICU due to the problem of RDS. Of the 1651 preterm newborns with RDS followed for 28 days, 436 (26.4%) died due to RDS, 383(23.2%) died due to other causes and 832(50.4%) survived and were discharged (discharged alive) from the centers. The distributions of the clinical characteristics of the neonates are presented in Table 4.1.

Of 1,651 enrolled neonates, 913 (55.3%) were males making the sex ratio 1.24:1. The highest proportion (39.6%) of the preterm infants in this study had a gestational age of 32 to 34 weeks. Among the 82 neonates with RDS having a gestational age of fewer than 28 weeks, 56.1% of them died due to RDS, 34.1 % died due to other causes and only 9.8 % were discharged alive from the hospitals. It was observed that the number of neonates being discharged alive from the hospital increases as gestational age increases. Out of the 139 preterm newborns with birth-weight less than 1Kg, 52.5% of them died due to RDS while 33.8% of them died due to other causes.

The mean and median follow up period of the study were about 9 and 6 days respectively. Moreover, the mean birth-weight of the preterm infants was 1.55kg and the median was 1.5 Kg. Similarly, the mean and median age of mothers who gave preterm birth was 26.3 and 26 years respectively. Details are shown in Table 4.2. Among the neonates included in the study, 393 (23.8%) were from SPH, 382(23%) from TAH, 326(19.7%) from JUH, 240(14.5%) from GMH and 210(12.7%) from GUH as shown in Table 4.3 of the appendix.

A graph of the absolute probability of the cumulative incidence of the event of interest and competing risks against follow up times is presented in Figure 4.2. Within the first 10 days of the follow-up time, neonates admitted to NICU had a higher probability of death due to RDS than death due to other causes or being discharged alive from the center. However, the probability of being discharged alive rose after this follow-up time and was higher than the probability of death throughout the follow-up period.

Table 4.1 Distribution of Clinical Characteristics of Neonates

			Health status of the neonate		
			Died due to RDS	Died due to Other Cause	Discharged alive
Covariates	**Categories**	**Total**	**Percent**	**Percent**	**Percent**
Sex	Female	738	27.5	22.5	50
	Male	913	25.5	23.8	50.7
Gestational Age (in weeks)	< 28	82	56.1	34.1	9.8
	28-31	648	37	30.1	32.9
	32-34	654	19.6	16.8	63.6
	>=35	267	8.2	18.7	73
Multiple pregnancies	No	1111	28.1	23.2	48.7
	Yes	540	23	23.1	53.9
Birth weight (in Kilo grams)	<1	139	52.5	33.8	13.7
	1-1.5	610	34.9	28.9	36.2
	1.5-2	570	24.1	18.6	60
	>=2	332	8.4	16.3	75.3
Maternal Age(in Years)	<20	130	27.7	23.1	49.2
	20-34	1367	26.0	22.9	51.1
	>=35	154	28.6	26.0	45.5
Pneumonia	No	1611	26.6	22.7	50.7
	Yes	40	20	42.5	37.5
Antenatal Care Received	No	111	32.4	32.4	35.1
	Yes	1540	26	22.5	51.5
Diabetes mellitus	No	1631	26.5	23.3	50.2
	Yes	20	20	15	65
C-section	No	974	27.4	24.8	47.7
	Yes	677	25	20.8	54.2
Cardiac disease	No	1637	26.4	23.3	50.3
	Yes	14	28.6	7.1	64.3
Feeding Problems	No	1265	24.3	21.9	53.8
	Yes	386	33.4	27.5	39.1
Hypertensive disorder	No	1173	25.2	24.3	50.5
	Yes	478	29.3	20.5	50.2

Figure 4.1: Plot of Cumulative incidence of the outcome Variables

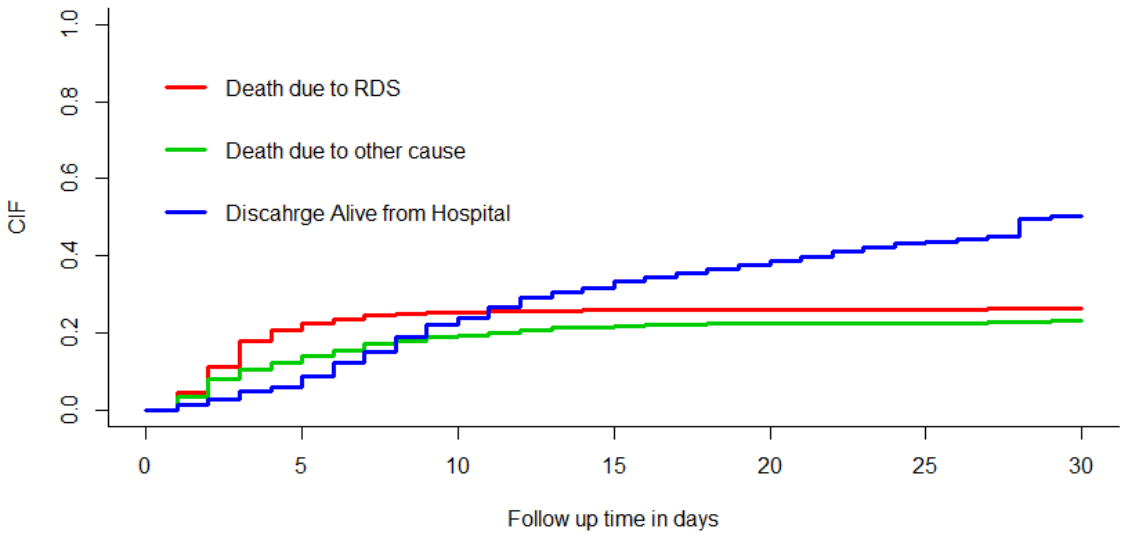


Table 4.2 Descriptive Statistics for Continuous Variables

Variables	Mean	SD	Median	Min	Max
Follow up time (Days)	8.91	7.93	6	1	30
Birth-weight (Kg)	1.55	0.46	1.5	0.4	3.4
Maternal Age(Years)	26.25	5.27	26	15	45
Gestational Age(Weeks)	31.76	2.51	32	22	36

4.2.Analytical Results

Gray's test (also called modified chi-square test) was used to assess the association between each potential prognostic factor and the outcome variables considered in the study.

The Gray's test results showed that meningitis, anemia, multiple pregnancies, gestational age, birth-weight, and feeding problem have a statistically significant association with the death of preterm infants due to RDS while pneumonia, meningitis, anemia, antenatal care received, gestational age, birth-weight, and feeding problem had a statistically significant association with the death of preterm infants due to other causes as presented in Table 4.4 below. Moreover, C-

section, antenatal care received, birth-weight, feeding problem, and gestational age have a significant association with discharge alive as shown in Table 4.5.

The competing risk data was analyzed using the **Cmprsk** package of the R statistical software (version 3.6.2). For multivariate analysis of the Fine-Gray Model, the **crr** function was used, and prognostic factors with P-value <0.05 were declared as significant.

Table 4.3. Gray's Test Results for death due to RDS and Other Causes

Covariates	**Categories**	**Death due to RDS**		**Death due to Other causes**	
		Gray's test	**P-value**	**Gray's test**	**P-value**
Pneumonia	No	0.95	0.33	8.33	0.004
	Yes				
Meningitis	No	4.54	0.03	16.1	0.00006
	Yes				
Anemia	No	11.0	0.0009	7.56	0.006
	Yes				
Multiple pregnancies	No	4.5	0.03	0.01	0.94
	Yes				
Sex	Male	0.52	0.47	0.55	0.46
	Female				
C-section	No	1.35	0.24	3.67	0.055
	Yes				
Diabetes mellitus	No	0.35	0.56	0.71	0.40
	Yes				
Birthweight (KG)	Less than 1	137.9	0.0000	35.8	0.0000
	1.0-1.5				
	1.5-2.0				
	2.0 or above				
Maternal age		29.62	0.43	35.4	0.19
Cardiac Disease	No	0.05	0.82	2.07	0.15
	Yes				
Antenatal Care Received	No	2.31	0.13	6.59	0.01
	Yes				
Hypertensive disorders	No	2.59	0.11	2.99	0.08
	Yes				
Gestational Age(in weeks)	Less than 28	145.3	0.0000	41.5	0.0000
	28-31				
	32-34				
	35 or above				
Feeding problem	No	12.7	0.00036	4.8	0.028
	Yes				

Table 4.4. Gray's Test Results for Discharged alive

Covariates	Categories	Survived to Discharge	
		Gray's test	P-value
Pneumonia	No	2.37	0.12
	Yes		
Meningitis	No	2.38	0.12
	Yes		
Anemia	No	0.89	0.344
	Yes		
Multiple pregnancy	No	2.75	0.097
	Yes		
Sex	Male	0.53	0.47
	Female		
C-section	No	6.15	0.01
	Yes		
Diabetes mellitus	No	1.39	0.24
	Yes		
Birthweight (KG)	Less than 1	247.7	<0.0001
	1.0-1.5		
	1.5-2.0		
	2.0 or above		
Maternal age		25.8	0.63
Cardiac Disease	No	0.41	0.52
	Yes		
Antenatal Care Received	No	6.51	0.01
	Yes		
Hypertensive disorders	No	0.37	0.54
	Yes		
Gestational Age(in weeks)	Less than 28	145.3	<0.0001
	28-31		
	32-34		
	35 or above		
Feeding problem	No	27.23	<0.0001
	Yes		

Figure 4.2 Plot of the Proportionality of the hazard of the CIF for Anemia

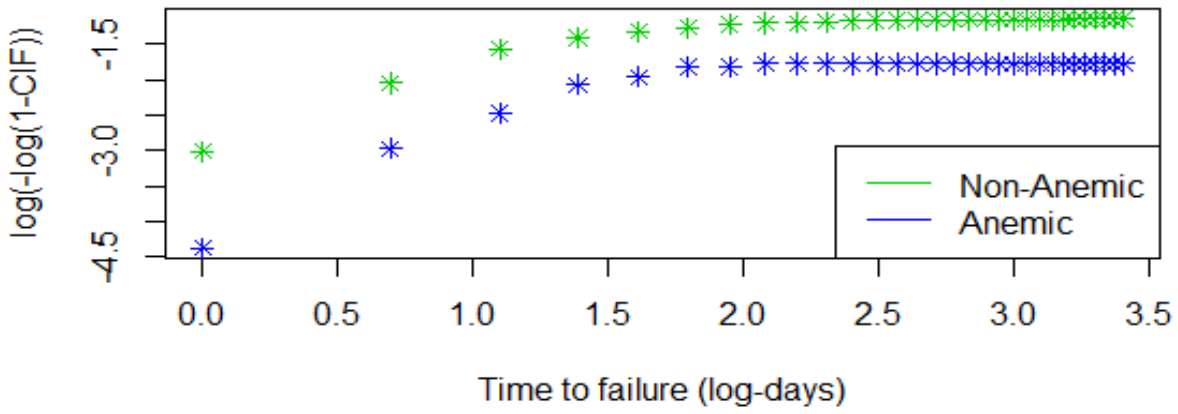
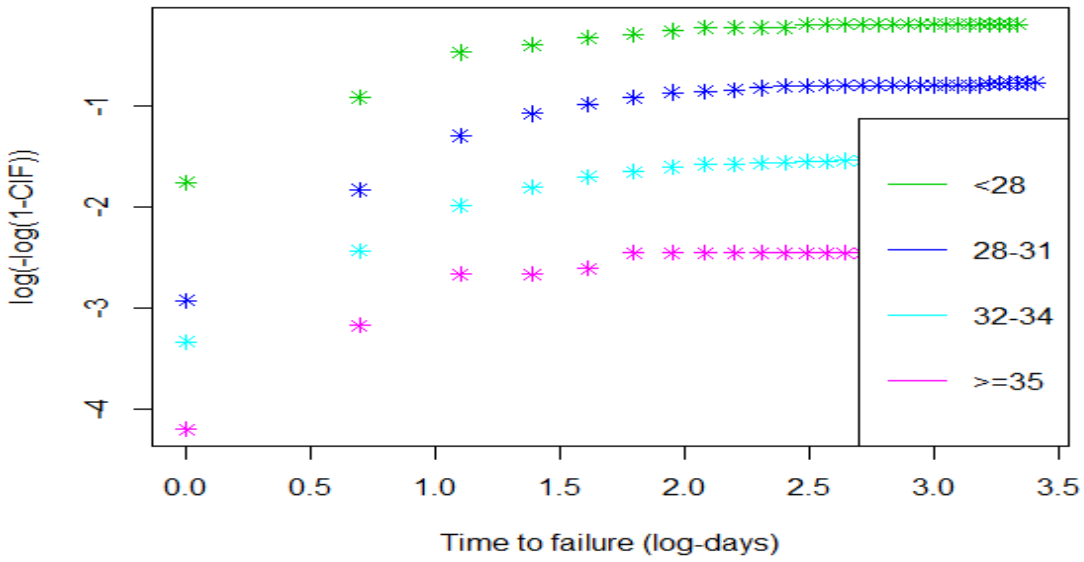


Figure 4.3: Plot of the Proportionality of the hazard of the CIF for Gestational Age



Sub-distribution hazard ratio (sHR) and 95% CI for sHR obtained from the Fine-Gray model fit results are presented in Table 4.5, Table 4.6, and Table 4.7.

The results show that Sex, C-section, diabetes mellitus, maternal age, cardiac disease, antenatal care received, hypertensive disorder and feeding problem have no significant influence on any causes of neonatal mortality. The preterm infants who had pneumonia are 93% more likely to die due to other causes but they have 40% less likely to be discharged alive than those preterm infants who did not have pneumonia. Anemic preterm infants were 60% less likely to die due to RDS whereas they were 36% more likely to die due to other causes compared to those without anemia. For neonates with birth weight 1000-1500, 1500-2000, and 2000g and above, the relative probabilities of failure or death due to RDS were 41%, 53%, and 76% less than for those with a birth weight of less than 1000g, respectively. While preterm infants with birth weight 1000-1500, 1500-2000, and 2000g and above increased the cumulative incidence of being discharged alive 2.06, 3.21, and 4.49 times than that for those with a birth weight of less than 1000g, respectively. Preterm infants with meningitis were 45% less likely to be discharged than those without meningitis.

Mothers with multiple pregnancies had a 22% lower risk of failure or death of neonates due to RDS than those mothers who have not experienced multiple pregnancies. Moreover, neonates with gestational age 28-31, 32-34, and 35 and above weeks had 21%, 50%, 70% lower risk of death due to RDS than preterm infants with gestational age less than 28 weeks, respectively. Conversely, neonates with gestational age 32-34 weeks had a 43% lower risk of death due to other causes than those with a gestational age of fewer than 28 weeks. Also, neonates with gestational age 28-31, 32-34, and 35 and above weeks were 2.39, 4.27, 5.0 times more likely to be discharged alive than preterm infants with gestational age less than 28 weeks, respectively.

The assumption of proportionality was checked and the plots do not indicate a violation of this assumption as shown in Figures 4.2 and 4.3. Furthermore, in health research decision about the importance of covariates depending on statistical significance is difficult. Since covariates that are not statistically significant may be practically significant (in a practical sense). As a result, we have reported the Fine-Gray model that contains both statistically significant and non-significant covariates.

4.5. Results of the Fine-Gray Model for Death due to RDS

		Death Due to RDS			
Covariates	**Categories**	**Coef**	**sHR**	**95% CI**	**P-value**
Pneumonia (Ref.No)	Yes	-0.21	0.81	0.42-1.56	0.53000
Meningitis (Ref.No)	Yes	-0.87	0.42	0.15-1.14	0.09000
Anemia (Ref.No)	Yes	-0.92	0.40	0.27-0.60	0.00001
Multiple pregnancy (Ref.No)	Yes	-0.25	0.78	0.63-0.96	0.02100
Sex (Ref.Female)	Male	0.058	1.06	0.88-1.14	0.54000
C-section (Ref.No)	Yes	-0.11	0.89	0.73-1.10	0.27000
Diabetes mellitus (Ref.No)	Yes	-0.29	0.75	0.30-1.87	0.53000
Birthweight (Ref. less than 1K g)	1.0-1.5	-0.54	0.59	0.46-0.75	0.00001
	1.5-2.0	-0.76	0.47	0.35-0.64	0.00002
	2.0 or above	-1.42	0.24	0.15-0.40	0.00000
Maternal Age (Ref. less than 20)	20-34	0.14	1.15	0.91-1.46	0.24000
	35 or above	0.16	1.17	0.76-1.80	0.48000
Cardiac disease (Ref.No)	Yes	0.04	1.04	0.38-2.86	0.94000
Antenatal care received (Ref.No)	Yes	-0.05	0.95	0.68-1.33	0.76000
Hypertensive disorders (Ref.No)	Yes	0.03	1.03	0.83-1.27	0.81000
Gestational Age (in weeks) (Ref.less than 28)	28-31	-0.24	0.79	0.56-1.10	0.16000
	32-34	-0.69	0.50	0.34-0.74	0.00060
	35 or above	-1.21	0.30	0.16-0.55	0.00001
Feeding problem (Ref.No)	Yes	0.06	1.07	0.84-1.35	0.60000

Where Ref denotes the reference category, Coef denotes the beta coefficients and sHR is equivalent to the exponential of the beta coefficient.

Table 4.6. Results of the Fine-Gray Model for death due to Other Causes

		Death due to Other Causes			
Covariates	**Categories**	**Coef**	**sHR**	**95% CI**	**p-value**
Pneumonia (Ref.No)	Yes	0.66	1.93	1.23-3.05	0.00450
Meningitis (Ref.No)	Yes	0.76	2.14	1.34-3.40	0.00140
Anemia (Ref.No)	Yes	0.31	1.36	1.00-1.84	0.04800
Multiple pregnancy (Ref.No)	Yes	0.02	1.02	0.82-1.28	0.85000
Sex (Ref.Female)	Male	0.17	1.18	0.97-1.45	0.10000
C-section (Ref.No)	Yes	-0.05	0.95	0.75-1.21	0.68000
Diabetes mellitus (Ref.No)	Yes	-0.29	0.75	0.24-2.33	0.62000
Birthweight (Ref. less than 1K g)	1.0-1.5	0.04	1.04	0.77-1.41	0.79000
	1.5-2.0	-0.30	0.74	0.50-1.10	0.13000
	2.0 or above	-0.40	0.67	0.41-1.10	0.11000
Maternal Age (Ref. less than 20)	20-34	-0.033	0.97	0.76-1.24	0.79000
	35 or above	0.06	1.06	0.64-1.76	0.83000
Cardiac disease (Ref.No)	Yes	-1.30	0.27	0.04-1.71	0.17000
Antenatal care received(Ref.No)	Yes	-0.32	0.72	0.51-1.03	0.07100
Hypertensive disorders (Ref.No)	Yes	-0.15	0.86	0.66-1.13	0.28000
Gestational Age(in weeks) (Ref.less than 28)	28-31	-0.18	0.89	0.58-1.37	0.59000
	32-34	-0.56	0.57	0.35-0.93	0.02400
	35 or above	-0.32	0.72	0.40-1.32	0.29000
Feeding problem(Ref.No)	Yes	0.15	1.16	0.87-1.54	0.31000

Table 4.7. Results of the Fine-Gray Model for Discharged alive

		Discharged alive from NICU			
Covariates	**Categories**	**Coef**	**sHR**	**95% CI**	**P-value**
Pneumonia (Ref.No)	Yes	-0.51	0.60	0.38-0.97	0.03600
Meningitis(Ref.No)	Yes	-0.60	0.55	0.34-0.91	0.02100
Anemia (Ref.No)	Yes	0.01	1.01	0.80-1.28	0.90000
Multiple pregnancy (Ref.No)	Yes	0.12	1.13	0.98-1.31	0.09800
Sex (Ref.Female)	Male	-0.13	0.88	0.76-1.02	0.07900
C-section (Ref.No)	Yes	0.09	1.09	0.94-1.28	0.25000
Diabetes mellitus(Ref.No)	Yes	0.20	1.23	0.78-1.93	0.38000
Birthweight (Ref. less than 1K g)	1.0-1.5	0.72	2.06	1.46-2.91	0.00004
	1.5-2.0	1.17	3.21	2.23-4.62	<0.0001
	2.0 or above	1.50	4.49	3.05-6.60	<0.0001
Maternal Age (Ref. less than 20)	20-34	-0.04	0.97	0.81-1.15	0.69000
	35 or above	-0.18	0.84	0.56-1.24	0.37000
Cardiac disease (Ref.No)	Yes	0.20	1.22	0.59-2.51	0.59000
Antenatal care received(Ref.No)	Yes	0.33	1.39	0.99-1.96	0.05500
Hypertensive disorders (Ref.No)	Yes	0.03	1.03	0.87-1.22	0.71000
Gestational Age(in weeks) (Ref.less than 28)	28-31	0.87	2.39	1.20-4.76	0.01300
	32-34	1.45	4.27	2.14-8.54	0.00004
	35 or above	1.61	5.00	2.45-10.22	0.00001
Feeding problem(Ref.No)	Yes	-0.16	0.85	0.69-1.06	0.14000

5. DISCUSSION

Respiratory distress syndrome has become one of the main health complications for preterm infants and it is considered to be the major cause of increased morbidity and mortality for neonates (Qaril et al., 2018). The study of neonatal data has to turn out to be one of the main research areas in developing countries like Ethiopia due to its paramount importance for the nation as a measure of achievement of global agenda like sustainable development goals (SDGs).

The current study indicated that 26.4% of the neonates admitted to the NICU died due to respiratory distress syndrome. This result is comparable with the findings by Swarnkar et al. (2015**)** that the mortality rate of preterm infants due to respiratory distress syndrome was 22.86%.

In the present study, only about one-fourth of the total preterm neonates had the event of interest. Had we used the standard survival analysis methods, about three-fourths of the study participants would have been treated as censored observations which would have led to statistical error such as the wrong magnitude of parameter estimates (over or underestimates) and incorrect conclusion. These limitations were handled using a competing risk model which allowed us to acknowledge the possible competing events. Therefore, the current study aimed to identify prognostic factors by accounting for the potential competing risks. In particular, the Fine-Gray competing risk model was used to analyze the prognostic factors associated with the health status of neonates with respiratory distress syndrome.

The result of this study showed that preterm infants having mothers with multiple pregnancy problems had a lower risk of dying due to RDS than neonates whose mothers did not experience multiple pregnancies. This result contradicts the findings by Baseer et al. (2020) that multiple gestation pregnancy was associated with a high risk of neonatal respiratory diseases.

The results of our study revealed that neonates with birth-weight 1000-1500, 1500-2000, and 2000g and above had increased cumulative incidence of being discharged alive compared to those with birth-weight less than 1000g. This result is consistent with the finding by Bahwal et al. (2020) that preterm infants with low birth weight had a higher risk of death due to RDS than those with normal weight. Similar findings were observed in the study done by Swarnkar et al. (2015) that the frequency of RDS is inversely related to gestational age and birth weight.

Moreover, the study by Aluvaala et al. (2019) showed that the probability of neonates that survived and were discharged alive from the NICU increases with the increase of birth weight. Qaril et al. (2018) have also observed in their study that premature infants/neonates with extremely low birth-weight (<1000g) had increased risk of death due to RDS.

The current study revealed that preterm infants with gestational age less than 28 weeks had a higher risk/probability of dying due to RDS than preterm infants with the gestational age 28-31, 32-34, or 34 and above weeks. Conversely, only preterm infants with gestational age 32-34 weeks had a lower risk/probability of dying due to other causes than preterm infants with gestational age less than 28 weeks. This result is consistent with findings by Saha et al. (2017) that mortality and morbidity of neonates were higher for preterm infants with low gestational age and low birth weight. Another study done in Iran had similar findings with the present study that infants with gestational age less than 25 weeks had increased neonatal mortality (Gargari et al., 2018).

The result of the present study confirmed that preterm infants with anemia had a lower risk of dying due to RDS but increased cumulative incidence of death due to other causes than neonates who do not have anemia. However, the findings by Hoxha et al. (2018) in Poland showed that anemic neonates had a higher incidence of RDS.

In the present study, it was observed that the probability of death of neonates due to RDS was higher in the first ten days than both deaths of neonates due to other causes and being discharged alive from the neonatal intensive care unit. This finding is consistent with the result of the study by Aluvaala et al. (2019) that the highest rate of neonatal mortality occurred in the first weeks of admission to the intensive care unit. Likewise, our result also revealed that pneumonic neonates have a higher risk of dying due to other causes and reduced relative risk/ probability of being discharged alive than non- pneumonic neonates.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1.Conclusions

The objective of the study was to examine and identify potential prognostic factors related to the health status of preterm infants with respiratory distress syndrome problem. The study was based on the secondary data obtained from the study of illness in the preterm project.

When employing survival analyses, it is essential to consider the presence of any competing risks. For unadjusted survival analysis, the Kaplan–Meier method can handle only one outcome and yield unreliable results for the estimation of survival probability in the presence of competing risks. Conversely, different approaches are available for multivariable survival analysis in a competing risk setting. Generally, the sub-distribution hazard is most appropriate for the prediction of a survival probability.

In our study, a competing risk modeling framework was applied. In particular, the Fine-Gray Model or sub-distribution hazard model was used to identify significant prognostic factors associated with the health status of neonates with RDS. Competing risk models are helpful to have separate parameter estimates for each recognized competing event.

In this study, the death of neonates due to other causes and neonates being discharged alive were events that compete with the event of interest, death of neonates due to RDS.

The study revealed that anemia, multiple pregnancies, birthweight, and gestational age are the prognostic factors significantly associated with the death of neonates due to RDS while pneumonia, meningitis, anemia, and gestational age of neonates were the significant prognostic factors related to the death of neonates due to other causes. Similarly, pneumonia meningitis, birthweight, and gestational age were identified as the significant prognostic factors associated with neonates being discharged alive.

6.2.Recommendations

This study recommends that health officials or practitioners should provide desired treatment giving more attention to neonates diagnosed for health complications of anemia, pneumonia, and meningitis to prolong their survival time in turn reduce neonatal mortality. Offering intensive and adequate treatments for those critically exposed neonates with the lowest birth-weight and

gestational age could decrease the burden of neonatal mortality. Moreover, evaluating the existing neonatal health care modalities provided in intensive care units may help to increase neonates' incidence of being discharged alive; which in turn reduces neonatal mortality in neonatal intensive care units. Further study with more prognostic factors is needed to see the concerns addressed in this study.

REFERENCES

Alfarwati TW, Alamri AA, Alshahrani MA, Al-Wassia H (2019). Incidence, Risk Factors, and Outcome of Respiratory Distress Syndrome in Term Infants at the Academic Centre, Jeddah, Saudi Arabia. *Med Arch. 2019 Jun;73(3) 183-186. DOI:10.5455/medarh.2019.73.183-186. PMID: 31402802; PMCID: PMC6643333.*

Altman M, Vanpée M, Cnattingius S, Norman M (2013). Risk factors for acute respiratory morbidity in moderately preterm infants. *Paediatr Perinat Epidemiol*. 2013 Mar;27(2):172-81. DOI: 10.1111/ppe.12035. PMID: 23374062.

Aluvaala J, Collins GS, Maina B, Mutinda C, Wayiego M, Berkley JA, English M (2019). Competing risk survival analysis of time to in-hospital death or discharge in a large urban neonatal unit in Kenya. *Wellcome Open Res*. 2019 Jun 17; 4:96. DOI: 10.12688/wellcomeopenres.15302.1. PMID: 31289756; PMCID: PMC6611136

Asmare Y, Asefaw H, Yirga T, Habtewold T, Sinshaw A, Shiferaw W. (2019). Incidence of Respiratory distress and its predictors among neonates admitted at neonatal intensive care unit, Black Lion Specialized Hospital, Addis Ababa, Ethiopia, 2019.DOI: 10.1101/19007823.

Austin PC, Fine JP (2017). Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat Med*. 2017Nov 30;36(27):4391-4400. DOI: 10.1002/sim.7501. Epub 2017 Sep 15. PMID: 28913837; PMCID: PMC5698744.

Austin PC, Lee DS, Fine JP (2016). Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016 Feb 9;133(6):601-609. DOI: 10.1161/CIRCULATIONAHA.115.017719. PMID: 26858290; PMCID: PMC4741409.

Bahwal SA, Jawass MA, Bahartha AS (2020)**.** Respiratory Distress and Its Outcome among Neonates Admitted to the Neonatal Intensive Care Unit of Mukalla Maternity and Child Hospital, Yemen. April 2020; 18 (1), DOI:https://DOI.org/10.26505/DJM.18014940915.

Baseer KAA, Mohamed M, Abd-Elmawgood EA (2020). Risk Factors of Respiratory Diseases Among Neonates in the Neonatal Intensive Care Unit of Qena University Hospital, Egypt. *Ann Glob Health*. 2020 Feb 26; 86(1): 22. DOI: 10.5334/aogh.2739. PMID: 32140431; PMCID: PMC7047767.

Condò V, Cipriani S, Colnaghi M, Bellù R, Zanini R, Bulfoni C, Parazzini F, Mosca F (2017). Neonatal respiratory distress syndrome: are risk factors the same in preterm and term

infants?, *The Journal of Maternal- Fetal & Neonatal Medicine*, 30:11, 1267-1272. DOI: 10.1080/14767058.2016.1210597.

Cox DR (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1972; *34*(2), 187-220. www.jstor.org/stable/2985181.

Fine JP, Gray RJ (1999). A Proportional Hazards Model for the Sub-distribution of a Competing Risk. *Journal of the American Statistical Association. 1999; 94*(446), 496-509. DOI:10.2307/2670170.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. Annals of Statistics, 16, 1141–1154.

Haller B (2014). The analysis of competing risks data with a focus on the estimation of cause-specific and sub-distribution hazard ratios from a mixture model. 2014. Corpus ID: 106561243.

Hinchliffe SR, Seaton SE, Lambert PC, Draper ES, Field DJ, Manktelow BN (2013). Modeling time to death or discharge in neonatal care: an application of competing risks. Paediatr Perinat Epidemiol. 2013 Jul; 27(4):426-33. DOI: 10.1111/ppe.12053.

Hubbard RM, Choudhury KM, Lim G (2018). Treatment Patterns and Clinical Outcomes in Neonates Diagnosed With Respiratory Distress Syndrome in a Low-Income Country: A Report From Bangladesh. *Anesth Analg.* 2018 May;126(5):1684-1686. DOI: 10.1213/ANE.0000000000002865. PMID: 29481437; PMCID: PMC5975634.

Kamath BD, Macguire ER, McClure EM, Goldenberg RL, Jobe AH (2011). Neonatal mortality from respiratory distress syndrome: lessons for low-resource countries. *Pediatrics*. 2011 Jun;127(6):1139-46. DOI: 10.1542/peds.2010-3212.

Kuk D, Varadhan R (2013). Model selection in competing risk regression. Stat Med. 2013 Aug 15;32(18):3077-88. DOI: 10.1002/sim.5762. Epub 2013 Feb 24. PMID: 23436643.

Leoce NM (2016). Prognostic Modeling in the Presence of Competing Risks: An Application to Cardiovascular and Cancer Mortality in Breast Cancer Survivors. https://DOI.org/10.7916/D89S1R41.

Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, Lawn JE, Cousens S, Mathers C, Black RE, (2016). Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet.* 2016; 388(10063): 3027–35. https://DOI.org/10.1016/S0140-6736(16)31593-8

Luo J, Yu J M, Chen J, Li Q, Feng Z (2019). Prevalence and Predictors of Infant Mortality of Respiratory Distress Syndrome in Beijing and Jiangxi Provinces of PR China: A Retrospective Cohort Analysis, *Iran Red Crescent Med J*; 2019; 21(6):e90679. DOI: 10.5812/ircmj.90679.

Mohammad KA, Fatima-Tuz-Zahura M, Bari W (2017). Fine and Gray competing risk regression model to study the cause-specific under-five child mortality in Bangladesh. *BMC Int Health Hum Rights* 17, 3 (2017). https://DOI.org/10.1186/s12914-017-0112-8.

Muhe, L. M., et al. (2019). Major causes of death in preterm infants in selected hospitals in Ethiopia (SIP): a prospective, cross-sectional, observational study. *The Lancet Global Health* **7**(8): e1130-e1138.

Niesłuchowska-Hoxha A, Cnota W, Czuba B, Ruci A, Ciaciura-Jarno M, Jagielska A, Wójtowicz D, Kierach R, Dąbrowski K, Sidorowicz M, Skrzypulec-Plinta W, Wloch A, Borowski D, Węgrzyn P (2018). A Retrospective Study on the Risk of Respiratory Distress Syndrome in Singleton Pregnancies with Preterm Premature Rupture of Membranes between 24+0 and 36+6 Weeks, Using Regression Analysis for Various Factors. Biomed Res Int. 2018 Oct 4;2018:7162478. DOI: 10.1155/2018/7162478. PMID: 30402491; PMCID: PMC6193337.

Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ (2013). When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant*. 2013 Nov;28(11):2670-7. DOI: 10.1093/ndt/gft355. Epub 2013 Aug 24. PMID: 23975843.

Orsido TT, Asseffa NA, Berheto TM (2019). Predictors of Neonatal mortality in the Neonatal intensive care unit at referral Hospital in Southern Ethiopia: a retrospective cohort study. BMC Pregnancy Childbirth. 2019 Feb 28;19(1):83. DOI: 10.1186/s12884-019-2227-5. PMID: 30819143; PMCID: PMC6396455.

Patil S, Halkude V, Tondare M, Mudaglimath A (2018). Study of Risk factors associated with the development of severe respiratory distress in the newborn. *Int J Contemp Pediatr* 2018; 5:2235-9. https://DOI.org/10.18203/2349-3291.ijcp20184287.

Pintilie M (2006). Competing Risks: A Practical Perspective. John Wiley & Sons: New York, 2006. pp 24.

Porta N, G´omez G, Calle ML, Malats N (2007). Competing Risks Methods.

Prentice, R. L., Kalbfeisch, J. D., Peterson, A. V., J., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, 34(4), 541-554.

Qaril SA, Alsufyani A A, Muathin SH, Margoushy NM (2018). Prevalence of Respiratory Distress Syndrome in Neonates**.***The Egyptian Journal of Hospital Medicine* (January 2018) Vol. 70 (2), Page 257-264.

Resche-Rigon M, Azoulay E, Chevret S (2005). Evaluating Mortality in Intensive Care Units: Contribution of Competing Risks Analyses. *Crit Care*. 2006 Feb;10(1): R5. DOI: 10.1186/cc3921. PMID: 16356211; PMCID: PMC1550827.

Sabzehei MK, Basiri B, Shokouhi M, Fayyazi A (2017). Causes and Outcomes of Respiratory Distress in Neonates Hospitalized in NICU of Be'sat Hospital in Hamadan Province, Iran. Int J Pediatr 2017; 5(12): 6253-60.DOI: 10.22038/ijp.2017.25776.2198

Saha L, Sarkar PK, Rahman M, Baidya M, Chowdhury MA (2017). The outcome of Preterm Neonates Presented with Respiratory Distress in a Tertiary Care Hospital of Bangladesh. *DS (Child) H J* . 2017; 33(1): 16-20.

Saleh Gargari S, Kashanian M, Zendedel H, Nayeri F, Shariat M, Haghollahi F (2018). Survival and Risk Factors of Extremely Preterm Babies (< 28 weeks) in the Three Iranian Hospitals. *Acta Med Iran*. 56(3):181-188. http://acta.tums.ac.ir/index.php/acta/article/view/6588.

Swarnkar K, swarnkar M (2015). Neonatal respiratory distress in the early neonatal period and its outcome. *International Journal of Biomedical and Advance Research*. 2015; 6(9), pp. 643-647. DOI: 10.7439/ijbar.v6i9.2430.

Tochie JN, Choukem SP, Langmia RN, Barla E, Koki-Ndombo P (2016). Neonatal respiratory distress in a reference neonatal unit in Cameroon: an analysis of prevalence, predictors, etiologies, and outcomes. *Pan Afr Med J*. 2016 Jun 21;24:152. DOI: 10.11604/pamj.2016.24.152.7066. PMID: 27642489; PMCID: PMC5012800.

UNICEF, Organization WH. Levels & trends in child mortality estimate developed by the UN inter-agency group for child mortality estimation report 2018.

Van Walraven C, McAlister FA (2016). Competing risk bias was common in Kaplan-Meier risk estimates published in prominent medical journals, *Journal of Clinical Epidemiology*, 2016 Jan;69:170-3.e8. DOI: 10.1016/j.jclinepi.2015.07.006.

Wang J, Liu X, Zhu T, Yan C (2015). Analysis of neonatal respiratory distress syndrome among different gestational segments. *Int J Clin Exp Med*. 2015 Sep 15;8(9)::16273-9. PMID: 26629144; PMCID: PMC4659032.

Wang, J, Yan J, Han J, Ning Y, Yan C (2019). Risk factors for respiratory distress syndrome among Chinese infants of 34-42 weeks gestational age: a multi-center observational study. *Int J Clin Exp Med*. 2019;12(4):4354-4360 www.ijcem.com_/ISSN:1940-5901/IJCEM0079039.

Wesenu M, Kulkarni S, Tilahun T (2017). Modeling Determinants of Time-To-Death in Premature Infants Admitted to Neonatal Intensive Care Unit in Jimma University Specialized Hospital. *Ann. Data. Sci.* 4**,** 361–381 (2017). https://DOI.org/10.1007/s40745-017-0107-2

Yismaw AE, Gelagay AA, Sisay MM (2019). Survival and predictors among preterm neonates admitted at the University of Gondar comprehensive specialized hospital neonatal intensive care unit, Northwest Ethiopia. *Ital JPediatr* 45**,** 4(2019).n https://DOI.org/10.1186/s13052-018-0597-3

Zhang Z. Survival analysis in the presence of competing risks (2017). Ann Transl Med 2017;5(3):47. DOI: 10.21037/atm.2016.08.62

APPENDIX

Table 4.8. Summary of Contribution of Hospitals to the Study with Outcome Variables

Selected Hospitals		Total	RDS	Other Causes	Discharged alive
Hospital name	GMH	240	21.7	13.3	65
	GUH	210	30.6	24.5	44.8
	JUH	326	20.2	24.8	54.9
	SPH	393	32.3	24.4	43.3
	TAH	382	25.1	25.7	49.2

Figure 4.4. The plot of Proportionality of the hazard of the CIF for Birthweight

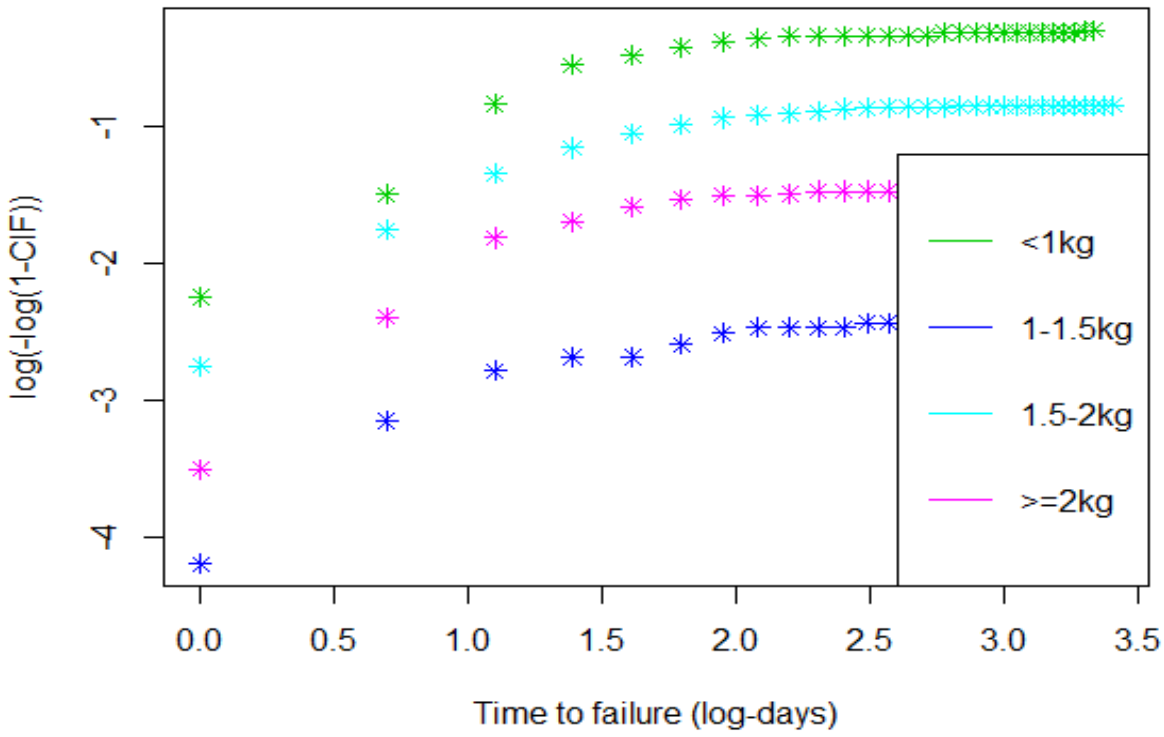


Figure 4.5. The plot of Proportionality of the hazard of the CIF for Multiple Pregnancy