



LUND
UNIVERSITY

FACULTY OF SCIENCE
CENTRE FOR MATHEMATICAL SCIENCES
DEPARTMENT OF MATHEMATICAL STATISTICS

MASTER THESIS

Infant Mortality Factors of Bangladesh

*A thesis presented in fulfillment of the requirements
for the degree of Master of Science*

in the

Department of Mathematical Statistics

Author :

Mohammad Sadek

Supervisor :

Anna Lindgren

June 6, 2016

Abstract

Infant mortality is a powerful indicator to measure the health and social status of a country. The purpose of this study is to identify the important factors that influence the infant mortality of Bangladesh. Moreover, this study suggests the government and the policy makers to improve the health and social sectors for reducing the country's infant mortality significantly. To investigate the important factors of infant (below 12 months) mortality of Bangladesh, data is used from the "Bangladesh Demographic and Health Survey (BDHS) 2011". Cross-tabulation, simple and multiple logistic regression techniques have been used to explore the predictors of infant mortality. Cross-tabulation and simple logistic regression analysis suggest that parental education, infant sex, vitamin A dose received, household toilet facility, number of antenatal visits, type of person giving care on mother before birth, birth order, tetanus injection before and after the pregnancy, birth type and type of person giving care on mother and infant after the birth have significant attention on infant mortality. For instances, risk of infant mortality 20%, 30% and 70% lower among the mothers having primary, secondary and higher education than those having no education respectively. Risk of infant mortality 52% and 81% lower among the mothers having tetanus injection than those having no tetanus injection before and after pregnancy respectively. On the other hand, multiple logistic regression analysis explore that father's education, infant sex, delivery by cesarean section, vitamin A dose received, place of delivery, type of person giving care on mother before birth, birth order, tetanus injection after the pregnancy, birth type and type of person giving care on infant are significant risk factors of infant mortality. Finally, these analyses have special attention on education for all, improved health services for infant and mothers to decrease the country's infant mortality.

Dedication

Dedicated to my parents, whose prayers and sacrifices lighted my way.

Declaration

I declare that this research paper submitted to the Lund university for the degree, Masters of Science (Mathematical Statistics) has not been submitted before by me at this institution or any other institution, but it is my own work and all contents in this paper has been fully acknowledged.

Mohammad Sadek

(Mohammad Sadek)

Date: 2016/06/06

Acknowledgements

I would express my immense indebtedness, sincere thanks and deep appreciation and regard to my reverend teacher and research supervisor, Anna Lindgren, for her untiring guidance, valuable suggestions, corrective advice and assistance for successful completion of this research work in a befitting manner.

I would also express my gratitude to Magnus Wiktorsson for offering me to study at Lund University. It never be forgotten for his continuous cooperation as a teacher and a program coordinator during the last two years.

Finally, I sincerely thank to my parents, family, and friends, who provide the advice and cooperation. This research would not be possible without all of them.

Contents

1	INTRODUCTION	9
1.1	Infant mortality in Bangladesh	9
1.2	Objective of the study	9
1.3	Source of data	10
1.4	Sampling	10
1.5	Selection of sample size	11
1.6	Handling with missing values of variables	11
1.7	Measurement of variables	11
2	GENERALIZED LINEAR MODEL	13
2.1	Introduction	13
2.2	Components of GLM	14
2.3	Assumptions of GLM	15
2.4	Multiple Logistic Regression:	15
2.5	Interpretation of the β coefficients in multiple logistic regression: . . .	17
2.5.1	Interpreting the intercept, β_0 :	17
2.5.2	Interpreting the slopes, β_j ; $j = 1, 2, \dots, p$:	17
2.6	Estimation of the β coefficients for a data set:	18
2.7	Assessing the goodness of fit of logistic regression model	18
3	INFANT MORTALITY ANALYSIS	20
4	RESULTS DISCUSSION AND CONCLUSION	33
4.1	Results Discussion	33
4.2	Conclusion	35
A	Bibliography	37

List of Figures

3.1	Administrative division wise infant mortality/1000 LBs	20
3.2	Infant mortality/1000 LBs at parents educational levels	21
3.3	Infant mortality/1000 LBs of different covariates	22
3.4	Infant mortality/1000 LBs of mother's age at first birth	23
3.5	Predicted probability of dying vs the observed response	24

List of Tables

1.1	Variables in infant mortality analysis.	12
2.1	Summary of GLMs	15
3.1	Infant mortality distribution	25
3.2	Relative risk of infant mortality against some socio-economic and demographic characteristics of the respondents	28
3.3	Nagelkerke and McFadden pseudo R^2 study.	30
3.4	Multiple logistic regression analysis of infant mortality	31

Chapter 1

INTRODUCTION

1.1 Infant mortality in Bangladesh

Bangladesh has an estimated 2015 population of 160.996 million which ranks 8th in the world. The country has a population density of 1033 people per square kilometer (United Nations World Population Prospects, 2015 Revision). Infant mortality is the death of an infant before his or her first birthday. The number of newborns dying under a year divided by the number of live births during a year defined as the rate of infant mortality. It is a powerful indicator on the nation's health because it is often associated with other health factors such as maternal health, quality and accessibility of medical care, and socioeconomic conditions. Infant mortality has long been very high in Bangladesh as compared to other countries in the region. According to BDHS reports, between 1990 and 2011, infant mortality rate fell less rapidly from 87/1000 to 44/1000 live births (LBs) over the last 20 years. The leading causes of infant mortality are maternal education, sanitation, access to safe drinking water, maternal and child health care services.

1.2 Objective of the study

This study has been made an attempt to identify the predictors of infant mortality and to suggest feasible strategies for the government to reduce the infant mortality in Bangladesh.

1.3 Source of data

This study utilizes the data extracted from the Bangladesh Demographic and Health Survey, 2011(BDHS-2011). BDHS-2011 was the sixth nationwide demographic and health survey in Bangladesh. The earlier BDHS were conducted in 1993-1994, 1996-97, 1999-2000, 2003-04, and 2007-08. All the ever married women aged 11-49 years who are the usual members of the selected households and among those who past the night before the survey in the selected households are eligible to be interviewed in the survey. The survey was designed in such a way that provides the representative results for the whole country, for the urban and rural areas separately and for seven administrative divisions of the country.

1.4 Sampling

The BDHS-2011 provides a sample which represents the country as a whole and its entire population as well. The survey used as a sample frame the list of enumeration areas (EAs) created by the Bangladesh Bureau of Statistics (BBS) for the fifth(conducted in 2011) Population and Housing Census of the People's Republic of Bangladesh. Each EA used as primary sampling unit (PSU) for the survey. The size of each EA on an average is 120 households. Administratively, Bangladesh is divided into seven divisions named Barisal, Chittagong, Dhaka, Khulna, Rajshahi, Rangpur, and Sylhet. Each of the administrative division is subdivided into *zilas* and *zila* into *upazilas*. Every *upazila* has urban and rural area. The uraban area in the upazila subdivided into *ward* and *ward* into *mohallas*. A rural area in the *upazila* subdivided int *union parishads* (UP) and *union parishads* into *mouzas*. The above zonal breaks up of the country also allow the country to separate into urban and rural areas.

The two-stage stratified sampling technique was employed to select the final sample. In the primary stage, 600 EAs were selected with probability proportional to the EA size (PPS), with 207 clusters in urban areas and 393 clusters in rural areas. A complete listing of all households is carried out in all selected EAs to provide a sampling frame for the second-stage selection of households. In the second-stage, a systematic sample of 30 households on average selected from each of 600 EAs to provide statistically reliable estimates of demographic, socio-economic and health characteristics for the country as a whole, for urban and rural areas separately, and for each of the seven administrative divisions also. According to design, the survey was conducted in 18,000 residential households, 6,210 in urban areas and 11,790 in rural areas. A total of 18,222 ever married women aged 11-49 years were identified in these households, and 17,842 were interviewed, yielding a response rate of 98 percent.

Data were collected from respondents during July 8, 2011 to December 27, 2011 in five different phases.

1.5 Selection of sample size

This study aims to investigate the infant mortality factors of Bangladesh based on the data from BDHS-2011. Therefore, this study requires a sub sample from the original sample. A sample of size 6878 is selected from the original sample of BDHS-211 for infant mortality analysis of Bangladesh where 6878 is the number of children who were born before 12 months of their guardian's interview date. Therefore, a sample of size 6878 is finally selected for the study.

1.6 Handling with missing values of variables

The dichotomous dependant variable has 6529 missing values. 41 is live and 308 is died after the first birthday of the children. We considered the 6529 individual as live kids. Finally our data becomes with 6570(41+6529) live kids and 308 dead kids. Any statistic based on this modification match with the statistic of the survey authority. For instance, the infant mortality calculated through this data is found 44 per 1000 live births which is found same in the publication of survey authority '*Bangladesh Demographic and Health Survey 2011*'.

Although I was interested to include some other variables in the analysis, it could not be possible to include them to the analysis due to a lot of missing values associated with them. For instances, the variables: Husband has other wives, Getting money needed for treatment, Distance-no nearby health facility etc.

1.7 Measurement of variables

The outcome variable of the study is a dichotomous variable produced by the respondents on the baby who was born before 12 months of their guardian's interview date. The respondents answered dichotomous as: i) the baby is alive and ii) the baby is died before the celebration of its first birthday.

There are 21 explanatory variables in this study. The outcome variable and the explanatory variables are given in the table 1.1.

Table 1.1: Variables in infant mortality analysis.

Variable	Description	Remark
dead	Life status of the baby after the first birthday	Factor: alive, died
MEd	Mother's education	Factor: No education, Primary, Secondary, Higher
FEd	Father's education	Factor: No education, Primary, Secondary, Higher
Res	Type of place of residence	Factor: Rural, Urban
MW	Mother's current work status	Factor: No, Yes
Sex	Sex of infant	Factor: Female, Male
FAge	Mother's age at the time of first birth	Factor: below 16, 16-20, 21 and above
Ces	Delivery by cesarean section	Factor: No, Yes
ViA	Vitamin A dose in first 2 months after delivery	Factor: No, Yes
FLi	Where the husband lives now	Factor: Living with her, Staying elsewhere
DWa	Drinking water for members of the household	Factor: not safe, safe
Toi	Toilet facility in the household	Factor: sanitary latrine, no sanitary latrine
PDel	Place of delivery	Factor: home, hospital/clinic
FW	Work status of father	Factor: currently employed, not currently employed
AntD	Number of antenatal visits during pregnancy	Factor: none, 1-2, 3-4, 5+
PreC	The type of person who gave parental care to the mother prior to the birth	Factor: no one, don't know, someone, doctor
BOrd	Birth order number	Factor: 1, 2-3, 4+
TetD	Tetanus injections during pregnancy	Factor: No, yes
TetB	Tetanus injections before pregnancy	Factor: No, yes
BTyp	Birth type	Factor: single, multiple
PCM	Person giving postnatal care on mother	Factor: medical personnel, others
PCK	Person giving postnatal check on infant	Factor: medical personnel, others

Chapter 2

GENERALIZED LINEAR MODEL

2.1 Introduction

In the classical linear regression, continuous response variable, $Y_i \sim N(\mu_i, \sigma^2)$. But for the binary response variable, Y_i , is far from the normality. In this case, the classical linear model is not suitable. Such a situation, the model has applied to some transformed value that is assumed to comply with the requirements of approximate conditional normality, linearity and constant variance. Many outcome measures in common use are not easy to successful transformation. In other cases, the original scale may be a natural one and effect estimates and inference are more easily understood on this scale. The generalized linear model (GLM) is a generalization of ordinary least squares regression. The GLM generalizes linear regression by allowing the linear model to be related to the response variable via a link function and by allowing the magnitude of the variance of each measurement to be a function of its predicted value.

Generalized linear model refers to a larger class of models popularized by McCullagh and Nelder (2nd edition, 1989). In these models, the response variable Y_i is assumed to follow an exponential family distribution with mean μ_i , which is assumed to be some function of $\mathbf{x}_i^T \boldsymbol{\beta}$. These models are described for non-normal responses. This allowed them to develop a general algorithm for maximum likelihood estimation in all of these models

In classical linear regression model, $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n)^T$ is the vector of observations of response variable, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ is the vector of

parameters, $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)^T$ is the vector of errors and $\mathbf{X}$ is a $n \times (p + 1)$ data matrix.

In regression set up, $\mathbf{X}_i = (1, \mathbf{X}_{1i}, \mathbf{X}_{2i}, \dots, \mathbf{X}_{pi})$ gives the levels or values of the covariates corresponding to the i -th response, $\mathbf{Y}_i$ ($i = 1, 2, \dots, n$). It is assumed that the errors are random variables with zero mean and constant variance, σ^2 . Further distributional assumption is that the errors are also normally distributed.

The generalized linear model, which incorporates the classical model described above as a special case, comprises of three components, namely the random component, the systematic component and a link function connecting the two.

2.2 Components of GLM

There are three components to any GLM:

- *Random Component* - refers that Y_i 's are independently distributed with $\mathbb{E}(\mathbf{Y}_i) = \mu_i$ and constant variance, σ^2 .
- *Systematic Component* - vector values of the covariates, $\mathbf{X}_i = (1, \mathbf{X}_{1i}, \mathbf{X}_{2i}, \dots, \mathbf{X}_{pi})$ produces a linear predictor $\boldsymbol{\eta}_i$ given by

$$\boldsymbol{\eta}_i = \sum_{j=0}^p \mathbf{X}_{ji} \boldsymbol{\beta}_j = \mathbf{X}_i \boldsymbol{\beta} \quad \text{where,} \quad \mathbf{X}_{0i} = 1$$

- *Link Function* - refers to the link between random and systematic components. It says how the expected value of the response relates to the linear predictor of explanatory variables; e.g., $\boldsymbol{\eta}_i = g(\mathbb{E}(\mathbf{Y}_i)) = \mathbb{E}(Y_i) = \mu_i$ for linear regression, this link is called identity function. , or $\boldsymbol{\eta}_i = \text{logit}(p_i)$ for logistic regression.

These models extend classical linear model in two ways:

- The random component may come from an exponential family rather than a normal family.
- The link function is some monotonous and differentiable function.

The table 2.1 provides a good summary of GLMs

Table 2.1: Summary of GLMs

Random distribution	Link nature	Link function	Inverse of link function
Normal	Identity	$\eta = \mu$	$\mu = \eta$
Gamma	Inverse	$\eta = \mu^{-1}$	$\mu = \eta^{-1}$
Inverse Gaussian	Inverse squared	$\eta = \mu^{-2}$	$\mu = \eta^{-1/2}$
Poisson	Log	$\eta = \ln(\mu)$	$\mu = e^\eta$
Binomial	Logit	$\eta = \ln \frac{\mu}{1-\mu}$	$\mu = \frac{e^\eta}{1+e^\eta}$

2.3 Assumptions of GLM

- The data $Y_1, Y_2, \dots, Y_n$ are independently distributed.
 - The dependent variable Y_i does not need to be normal, but it assumes a distribution from an exponential family (e.g. Binomial, Poisson, Multinomial etc.)
 - GLM requires a linear relationship between the endogenous variable and the exogenous variables, but it assumes a linear relationship between the transformed response in terms of the link function and the exogenous variables.
 - Independent (explanatory) variables can be even the power terms or some other nonlinear transformations of the original independent variables.
 - The homogeneity of variance does not need to be satisfied.
 - Errors are independent but not normal.
 - GLM uses maximum likelihood estimation (MLE) instead of ordinary least squares (OLS) to estimate the parameters, and thus depends on large-sample approximations.
 - Goodness-of-fit measures depend on sufficiently large samples, more than 80% of the expected cells counts are greater or equal to 5.

2.4 Multiple Logistic Regression:

Logistic regression analysis is a widely used and popular analysis that is similar to linear regression analysis exception is that the outcome is dichotomous (e.g., died/lived

or yes/no or success/failure). In Logistic regression analysis, we calculate the odds of an outcome, and by using the natural log of the odds of the outcome as the dependent variable, the relationships can be linearized and treated much similar as multiple linear regression.

Simple logistic regression analysis is the application of regression with one dichotomous outcome and one independent variable; on the other hand, in multiple logistic regression analysis, we apply single dichotomous outcome like as simple logistic regression but more than one independent variable.

The dichotomous outcome in logistic regression is usually coded as 0 or 1, where 0 indicates that the outcome of interest is absent, and 1 indicates that the outcome of interest is present. If we say $\pi(X)$ represents the probability of an event that depends on p independent variables. Then, using an inverse logit formulation for modeling the probability, we have:

$$\pi(\mathbf{X}) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}} \quad (2.1)$$

To get the appropriate logit function from this, we calculate ($\mathbf{X} = (X_0, X_1, X_2, \dots, X_p)$; $X_0 = 1$):

$$\begin{aligned} \text{logit}[\pi(\mathbf{X})] &= \ln \left[\frac{\pi(\mathbf{X})}{1 - \pi(\mathbf{X})} \right] \\ &= \ln \left[\frac{\frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}}{1 - \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}} \right] \\ &= \ln \left[\frac{\frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}}{\frac{1}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}} \right] \\ &= \ln [e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}] \\ &= \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \\ &= \mathbf{X}\boldsymbol{\beta}; \text{ where } \boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p) \end{aligned} \quad (2.2)$$

Equation 2.1 provides the probabilities of outcome events given the covariate values $\mathbf{X} = (X_0, X_1, X_2, \dots, X_p)$ and equation 2.2 shows how the logistic regression is a ordinary linear regression model, once we transform the dichotomous outcome by the logit transformation. This transformation changes the range of $\pi(\mathbf{X})$ from 0 to 1 to $-\infty$ to $+\infty$, as usual for linear regression.

2.5 Interpretation of the β coefficients in multiple logistic regression:

2.5.1 Interpreting the intercept, β_0 :

Setting all the covariate values to zero in equation 2.1 and 2.2, then we get the following equation 2.3 and 2.4 respectively.

$$\pi(X_0) = \frac{e^{\beta_0}}{1 + e^{\beta_0}} \quad (2.3)$$

$$\beta_0 = \ln \left[\frac{\pi(X_0)}{1 - \pi(X_0)} \right] = \ln[Odds(\pi(X_0))] \quad (2.4)$$

Therefore β_0 sets the “baseline” event rate, through the above function in 2.3 and 2.4, when all covariate values are set to zero.

From equation 2.3 it is obvious that $\pi(X_0)$ is greater than 0.5, less than 0.5 and equal to 0.5 for $\beta_0 = \text{positive, negative}$ and 0 respectively, when all covariates are set to zero.

2.5.2 Interpreting the slopes, β_j ; $j = 1, 2, \dots, p$:

Setting all the covariate values to zero in equation 2.2 except X_j , then we get the following equation 2.5.

$$\begin{aligned} \ln \left[\frac{\pi(X_j)}{1 - \pi(X_j)} \right] &= \beta_0 + \beta_j = \ln \left[\frac{\pi(X_0)}{1 - \pi(X_0)} \right] + \beta_j \\ \implies \beta_j &= \ln \left[\frac{\pi(X_j)}{1 - \pi(X_j)} \right] - \ln \left[\frac{\pi(X_0)}{1 - \pi(X_0)} \right] = \ln \left[\frac{\frac{\pi(X_j)}{1 - \pi(X_j)}}{\frac{\pi(X_0)}{1 - \pi(X_0)}} \right] = \ln(OR_j) \\ \implies e^{\beta_j} &= OR_j \end{aligned} \quad (2.5)$$

Therefore, we see that the coefficient β_j is such that e^{β_j} is the ratio of $Odds(\pi(X_j))$ to $Odds(\pi(X_0))$ for a unit change in X_j .

2.6 Estimation of the β coefficients for a data set:

The dichotomous dependent variable associated with logistic regression is distributed as Binomial. For n observations with covariate values $\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{pi})$ $i = 1, 2, \dots, n$, the likelihood function is:

$$\prod_{i=1}^n \pi(\mathbf{X}_i)^{z_i} (1 - \pi(\mathbf{X}_i))^{1-z_i} \quad (2.6)$$

To estimate the unknown β coefficients, we have to maximize this likelihood function. We use the usual steps, first taking the logarithm of the likelihood function, and then taking $(p + 1)$ partial derivatives with respect to each β_j ; $j = 0, 1, 2, \dots, p$ and setting these $(p + 1)$ equations equal to zero, to form a set of $(p + 1)$ equations in $(p + 1)$ unknowns. Finally, solving this system of equations, we get the ML estimate of β_j 's.

We omit the elaborate calculation as we get the maximum likelihood estimates of β_j 's using statistical software.

Inferences also depend on SE formulae for confidence intervals, and likelihood ratio test for hypothesis testings. we will omit the detail calculation, and we rely on statistical software.

2.7 Assessing the goodness of fit of logistic regression model

The dichotomous response variables are analysed with the use of logistic regression model. Once data is fitted to the given data set, our next step is to investigate how good the considered model to the observed data set. In the context of logistic regression analysis, goodness of fit statistic designed to determine the adequacy or inadequacy of the fitted model in describing the relationship between response variable and potential risk factors. The main purpose of goodness of fit is to identify the model that best fits the data, otherwise interpretations may be incorrect or misleading. Otherwise, we have to search a suitable model, one that will be more useful in explaining the response variable.

Although several goodness of fit statistics have been suggested over the last forty years, none of them are to be uniquely best statistic in assessing the suitability of the logistic regression model. Among them, one of the popular goodness of fit statistics,

Nagelkerke R^2 , is noted in equation 2.7.

$$R_{Nagelkerke}^2 = \frac{1 - \left(\frac{L_0}{L_1}\right)^{2/n}}{1 - L_0^{2/n}} \quad (2.7)$$

With

L_0 = likelihood function under the null model

L_1 = likelihood function under the full model

$R_{Nagelkerke}^2 \in [0, 1]$

When the full model explain the outcome perfectly and has a likelihood of 1, then $R_{Nagelkerke}^2 = 1$ and when $L_0=L_1$, then $R_{Nagelkerke}^2$ equal to 0.

The other popular goodness of fit statistic is McFadden's pseudo R^2 is shown in the equation 2.8.

$$R_{McF}^2 = 1 - \frac{\ln L_1}{\ln L_0} \quad (2.8)$$

With

$\ln L_0$ = Log likelihood function under the null model

$\ln L_1$ = Log likelihood function under the full model

Chapter 3

INFANT MORTALITY ANALYSIS

Infant mortality/1000 Live Births (LBs) of different administrative division is shown in figure 3.1.

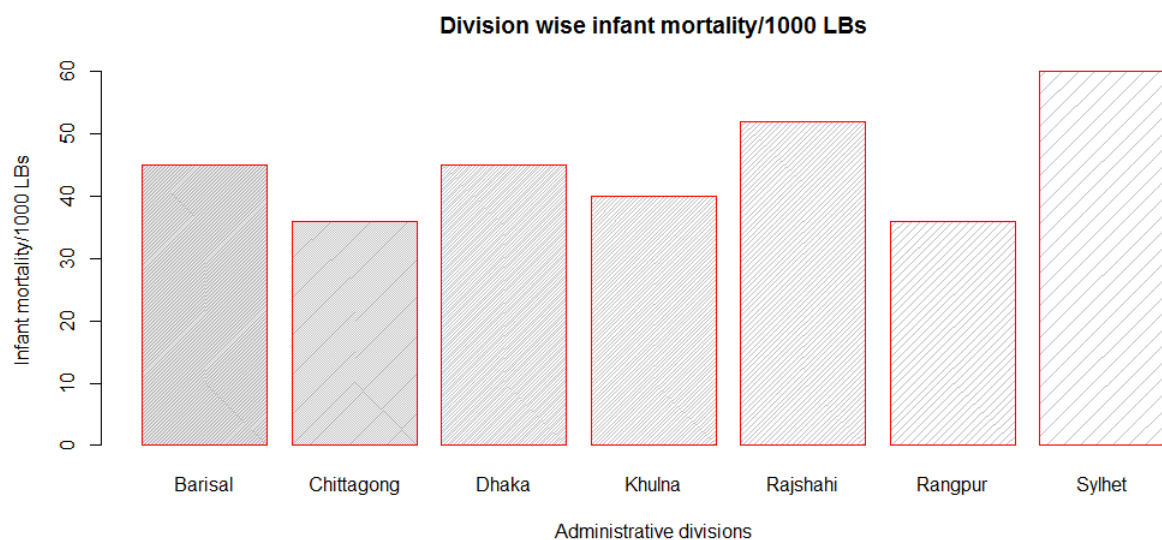


Figure 3.1: Administrative division wise infant mortality/1000 LBs

The bar chart in figure 3.1 shows how second infant mortality varies among seven administrative divisions of Bangladesh. It is clearly evident that the lowest infant mortality belongs to Chittagong and Rangpur each with 36/1000LBs. However, the highest infant mortality, 6/1000LBs, is observed in Sylhet. If we look at other four

divisions (Khulna, Barisal, Dhaka and Rajshahi) we can see that the infant mortality varies between 40-52/1000LBs.

Infant mortality/1000 Live Births (LBs) at parents educational levels is shown in figure 3.2.

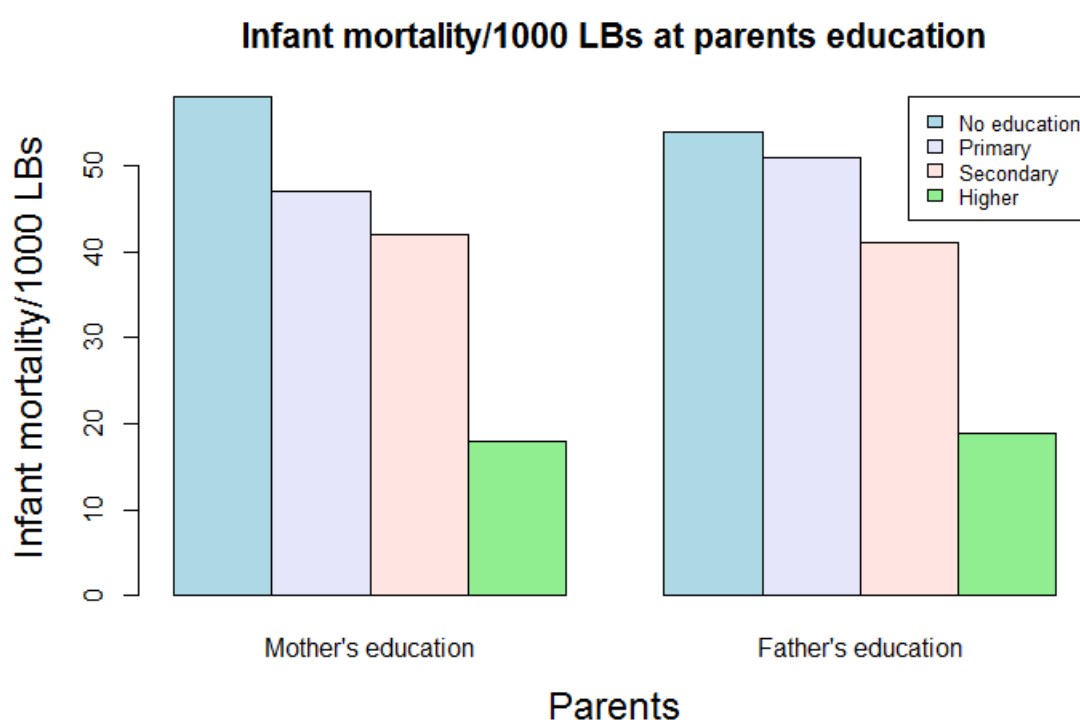


Figure 3.2: Infant mortality/1000 LBs at parents educational levels

The bar chart in figure 3.2 illustrates the the infant mortality at educational levels between parents. It can be seen that for both the parents, infant mortality decreases as their educational level increases.

Infant mortality/1000 Live Births (LBs) of different covariates is shown in figure 3.3.

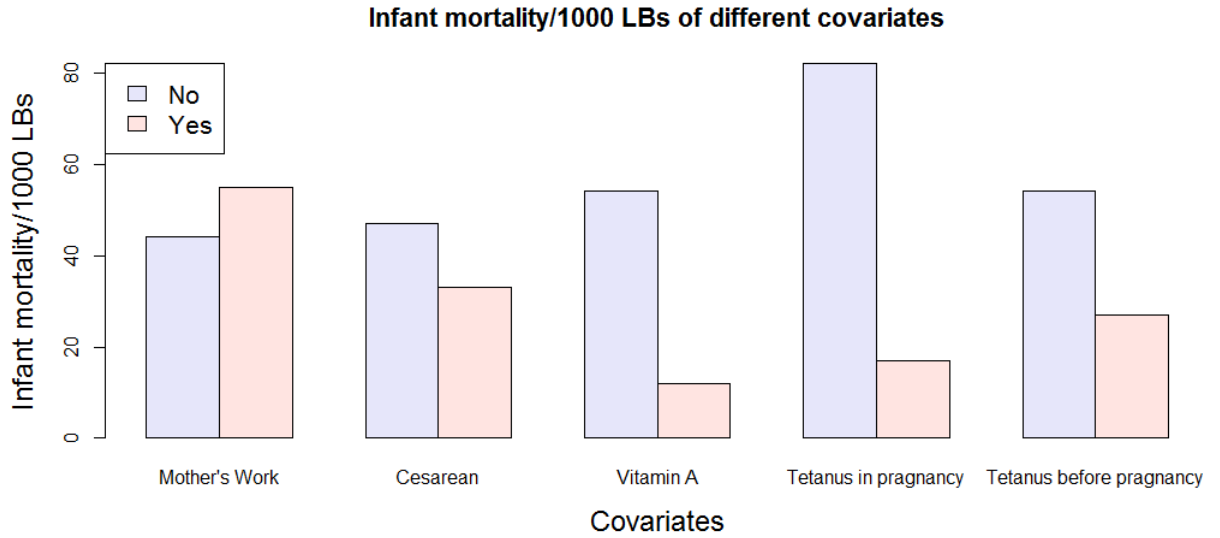


Figure 3.3: Infant mortality/1000 LBs of different covariates

The bar chart in figure 3.3 described the infant mortality over the five covariates each with two levels. It can be seen that is higher for the employed mothers than those who are not employed. It also can be seen that infant mortality for cesarean section is lower than the non-cesarean section. Another noticeable thing is that infant mortality is one-fifth for the babies who received vitamin A than those who did not receive vitamin A within the two months after the birth. The infant mortality is also one-fifth for the mothers who received tetanus injection during the pregnancy than the those who did not receive tetanus injection during the pregnancy. Further, infant mortality is almost half for the mothers who received tetanus injection before the pregnancy than those who did not receive tetanus injection before the pregnancy.

Infant mortality/1000 Live Births (LBs) of mother's age at first birth is shown in figure 3.4.

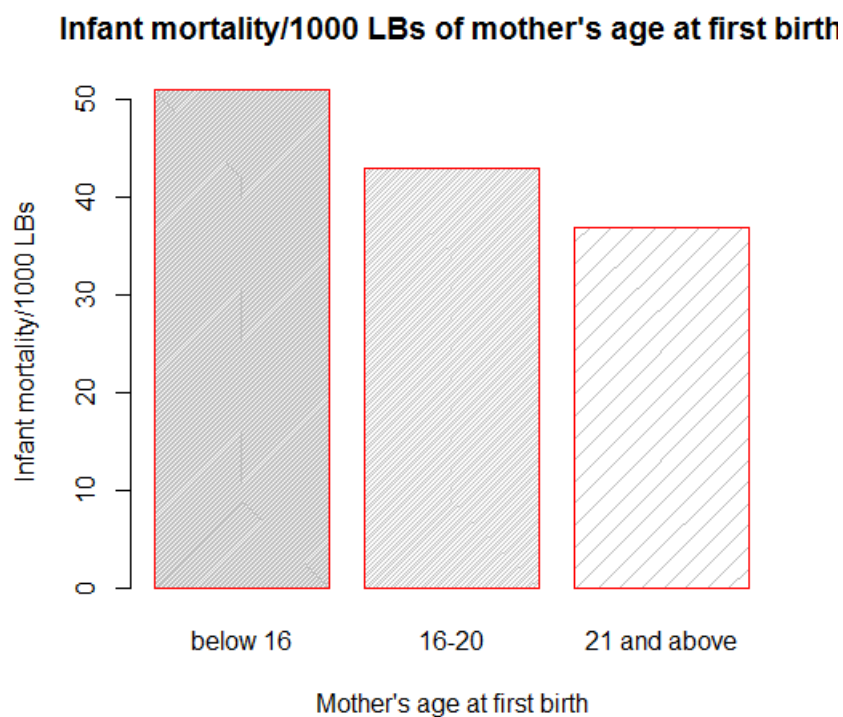


Figure 3.4: Infant mortality/1000 LBs of mother's age at first birth

This bar chart shows how infant mortality varies mother's age at first birth. It is immediately apparent that highest infant mortality occurs for the mothers who gave the first birth at below 16 years and lowest infant mortality occurs for the mothers who gave the first birth at 21 years and above. The infant mortality for the mothers who give the first birth at 16-20 years lies between the infant mortality for the mother's age at first birth below 16 years 21 and above years.

A plot of predicted probability of dying vs the observed response is shown in figure 3.5 on the basis of fitted multiple logistic regression.

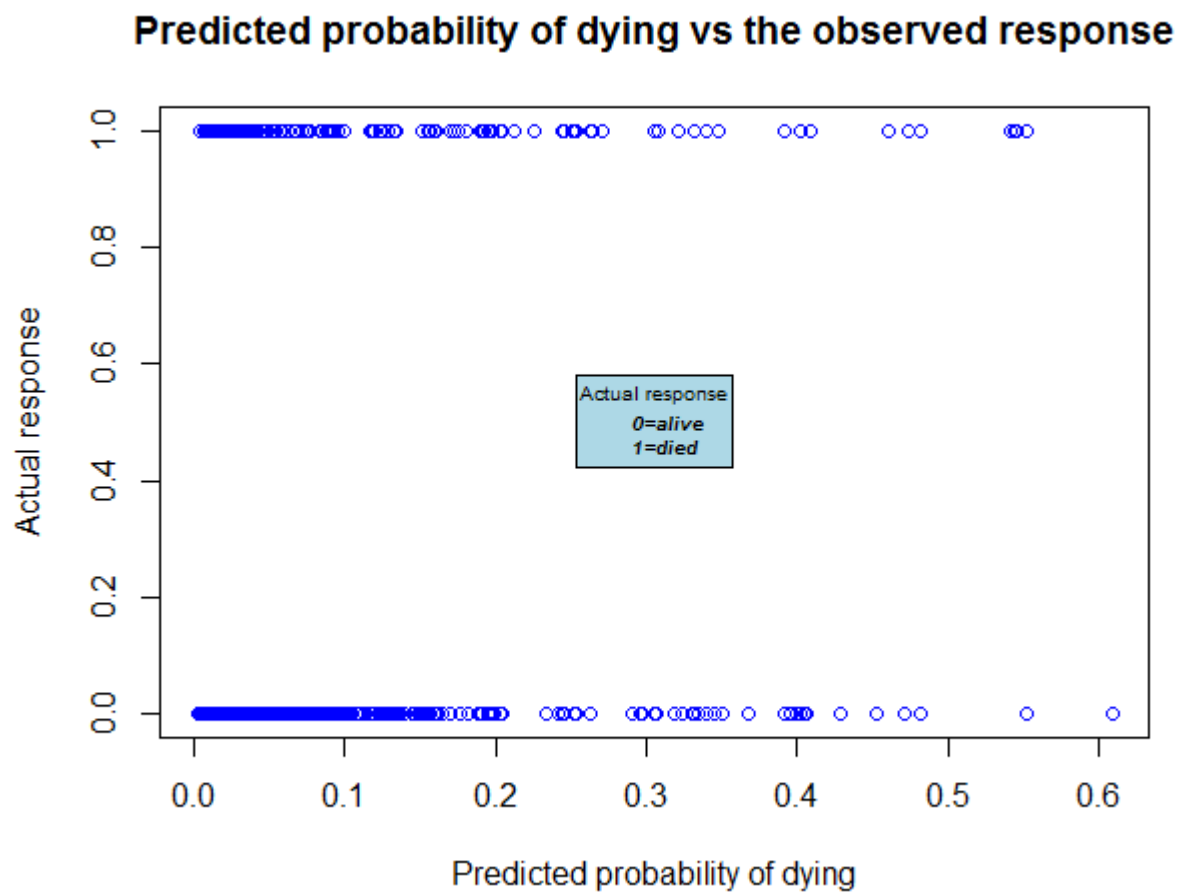


Figure 3.5: Predicted probability of dying vs the observed response

Table 3.1: Infant mortality distribution

Selected Variables	Number of infants dead	Total number of infants	Percent of death	p-value
Mother's education				
No education	81	1391	5.82	0.001241
Primary	101	2147	4.70	
Secondary	116	2793	4.15	
Higher	10	547	1.83	
Father's education				
No education	109	2003	5.44	9.614e-05
Primary	102	2001	5.10	
Secondary	79	1937	4.08	
Higher	18	937	1.92	
Type of place of residence				
Rural	212	4760	4.45	0.9341
Urban	96	2118	4.53	
Mother's current work status				
No	266	6111	4.35	0.1852
Yes	42	767	5.48	
Sex of infant				
Female	130	3360	3.87	0.0199
Male	178	3518	5.06	
Mother's age at the time of first birth				
below 16	123	2432	5.06	0.1792
16-20	155	3636	4.26	
21 and above	30	810	3.70	
Delivery by cesarean section				
No	277	5939	4.66	0.07326
Yes	31	939	3.30	
Vitamin A dose in first 2 months after delivery				
No	291	5441	5.35	1.836e-11

Continued on next page

Table 3.1 – *Continued from previous page*

Selected Variables	Number of infants dead	Total number of infants	Percent of death	p-value
Yes	17	1437	1.18	
Where the husband lives now				0.3863
Living with her	269	5879	4.58	
Staying elsewhere	39	999	3.90	
Drinking water for members of the household				0.08968
not safe	41	707	5.80	
safe	267	6171	4.33	
Toilet facility in the household				0.02782
sanitary latrine	123	3178	3.87	
no sanitary latrine	185	3700	5.00	
Place of delivery				0.6053
hospital/clinic	83	1756	4.73	
home	225	5122	4.39	
Work status of father				0.847
currently employed	303	6745	4.49	
not currently employed	5	133	3.76	
Number of antenatal visits during pregnancy				2.2e-16
none	247	3331	7.42	
1-2	28	1475	1.90	
3-4	15	1104	1.36	
5+	18	968	1.86	
Type of person giving prenatal care to the mother				2.2e-16
doctor	42	2236	1.88	
don't know	185	1408	13.14	
no one	62	1916	3.24	
someone	19	1318	1.44	
Birth order number				0.01711
1	130	2402	5.41	
2-3	121	3166	3.82	

Continued on next page

Table 3.1 – *Continued from previous page*

Selected Variables	Number of infants dead	Total number of infants	Percent of death	p-value
4+	57	1310	4.35	
Tetanus injections during pregnancy				2.2e-16
No	242	2948	8.21	
Yes	66	3930	1.68	
Tetanus injections before pregnancy				2.029e-07
No	244	4490	5.43	
Yes	64	2388	2.68	
Single or multiple births				2.2e-16
multiple	30	120	25.00	
single	278	6758	4.11	
Person giving postnatal care on mother				1.047e-07
Medical personnel	44	1868	2.36	
others	264	5010	5.27	
Person giving postnatal care on infant				5.763e-08
Medical personnel	50	2060	2.43	
others	258	4818	5.35	

Table 3.2: Relative risk of infant mortality against some socio-economic and demographic characteristics of the respondents

Risk factors	Relative risk of mortality	Relative risk of mortality (% lower w.r.t. <i>RC</i>)
Mother's education		
No education (<i>RC</i>)	1	
Primary	0.79836598	20
Secondary	0.70080291*	30
Higher	0.30117020***	70
Father's education		
No education (<i>RC</i>)	1	
Primary	0.93331594	7
Secondary	0.73881356*	26
Higher	0.34033803***	66
Mother's current work status		
No (<i>RC</i>)	1	
Yes	1.27295826	27 [‡]
Sex of infant		
Female (<i>RC</i>)	1	
Male	1.32413634*	32 [‡]
Delivery by cesarean section		
No (<i>RC</i>)	1	
Yes	0.69785620	30
Vitamin A dose in first 2 months after delivery		
No (<i>RC</i>)	1	
Yes	0.21187261***	79
Toilet facility in the household		
no sanitary latrine (<i>RC</i>)	1	
sanitary latrine	0.76497545*	24
Place of delivery		
home (<i>RC</i>)	1	
hospital/clinic	1.0797662	8 [‡]
Number of antenatal visits during pregnancy		

Continued on next page

Table 3.2 – Continued from previous page

Risk factors	Relative risk of mortality	Relative risk of mortality (% lower w.r.t. <i>RC</i>)
none (<i>RC</i>)	1	
1-2	0.24160556***	76
3-4	0.17198113***	83
5+	0.23657362***	76
Type of person giving prenatal care to the mother		
no one (<i>RC</i>)	1	
don't know	4.52338248***	352‡
someone	0.43738381**	56
doctor	0.57244098**	43
Birth order number		
1 (<i>RC</i>)	1	
2-3	0.69448528**	31
4+	0.79503960	20
Tetanus injections during pregnancy		
No (<i>RC</i>)	1	
Yes	0.19099380***	81
Tetanus injections before pregnancy		
No (<i>RC</i>)	1	
Yes	0.47921898***	52
Single or multiple births		
single (<i>RC</i>)	1	
multiple	7.76978417***	677‡
Person giving postnatal care on mother		
others (<i>RC</i>)	1	
Medical personnel	0.43366228***	57
Person giving postnatal check on infant		
others (<i>RC</i>)	1	
Medical personnel	0.43966215***	56

RC-Reference category, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ‡ % higher w.r.t. *RC*

Nagelkerke and McFadden pseudo R^2 study to get the better model is shown in table 3.3.

Table 3.3: Nagelkerke and McFadden pseudo R^2 study.

Model	R_N^2	R_{McF}^2
Use all predictors	0.174	0.150
Exclude MEd	0.173	0.149
Exclude Res	0.173	0.149
Exclude FAge	0.174	0.149
Exclude FLi	0.174	0.150
Exclude DWa	0.174	0.149
Exclude Toi	0.174	0.150
Exclude FW	0.174	0.150
Exclude AntD	0.174	0.149
Exclude TetB	0.174	0.150
Exclude PCM	0.174	0.150
Exclude PCK	0.174	0.149
Exclude FEd	0.169	0.146
Exclude MW	0.171	0.147
Exclude Sex	0.171	0.147
Exclude Ces	0.173	0.148
Exclude ViA	0.170	0.146
Exclude PDel	0.171	0.147
Exclude PreC	0.168	0.144
Exclude BOrd	0.172	0.148
Exclude TetD	0.169	0.145
Exclude BTyp	0.158	0.136
Use FEd, MW, Sex, Ces, ViA, PDel, PreC, BOrd, TetD, BTyp	0.168	0.145
Use MEd, Res, FLi, DWa, Toi, FAge, FW, AntD, TetB, PCM, PCK	0.076	0.065

Model with all the predictors, the value of Nagelkerke and McFadden pseudo R^2 are around 0.17 and 0.15 respectively. Dropping a significant predictor results in a notable decrease in the values of pseudo R^2 s, while no significant change occurs if we drop the insignificant predictors. R_N^2 and R_{McF}^2 are near zero when we exclude all the significant predictors.

Table 3.4: Multiple logistic regression analysis of infant mortality

Selected Characteristics	Estimate	<i>S.E</i>	Odds & OR	95% CI of OR
(Intercept)	-2.889	0.231	0.0557	0.0350 - 0.0868
Father's education				
No education			ref.	
Primary	-0.012***	0.150	0.9882	0.7361 - 1.3255
Secondary	-0.294	0.169	0.7454	0.5334 - 1.0373
Higher	-1.130***	0.288	0.3229	0.1786 - 0.5556
Mother's current work status				
No			ref.	
Yes	0.547**	0.181	1.7279	1.1979 - 2.4398
Sex of infant				
Female			ref.	
Male	0.327**	0.123	1.3868	1.0913 - 1.7668
Delivery by cesarean section				
No			ref.	
Yes	-0.499*	0.248	0.6074	0.3701 - 0.9834
Vitamin A				
No			ref.	
Yes	-0.789**	0.269	0.4544	0.2592 - 0.7500
Place of Delivery				
home			ref.	
hospital/clinic	0.592**	0.189	1.8084	1.2410 - 2.6055
Type of person giving prenatal care to the mother				
no one			ref.	
don't know	0.893**	0.196	2.4429	1.6763 - 3.6148
someone	-0.889**	0.273	0.4112	0.2347 - 0.6895
doctor	-0.500*	0.229	0.6063	0.3845 - 0.9465
Birth order number				
1			ref.	
2-3	-0.280*	0.140	0.7561	0.5744 - 0.9947

Continued on next page

Table 3.4 – *Continued from previous page*

Selected Characteristics	Estimate	<i>S.E</i>	Odds & OR	95% CI of OR
4+	-0.319	0.184	0.7266	0.5040 - 1.0372
Tetanus injection during pregnancy				
No			ref.	
Yes	-0.798***	0.186	0.4503	0.3130 - 0.6497
Birth type				
single			ref.	
multiple	1.578***	0.247	4.8445	2.9559 - 7.7908
Person giving postnatal care on kids				
other			ref.	
Medical personnel	0.372	0.229	1.4505	0.9219 - 2.2621

OR-Odds Ratio, **CI**-Confidence interval, ref.-reference category

*- $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 4

RESULTS DISCUSSION AND CONCLUSION

4.1 Results Discussion

The results obtained from infant mortality distribution presented in table 3.1. The result suggest that infant mortality risk is strongly related to parental education. The infant mortality are 5.82%, 4.70%, 4.15%, 1.83% for the group of mothers with the education illiterate, primary, secondary and higher respectively. It also be observed that infant mortality are 5.44%, 5.10%, 4.08%, 1.92% for the group of fathers with the education illiterate, primary, secondary and higher respectively. So, it can be concluded that the risk of infant mortality decreases as the parental educational level increases. Infant mortality are 4.45% and 4.53% in the rural and urban area respectively but from the result it is evident that residency status has no significant effect on infant mortality. Although the infant mortality rate for the cesarean and non-cesarean infant are 4.66% and 3.30% respectively, the result support that delivery by cesarean and non-cesarean section has no significant effect on infant mortality. The other factors that have found no significant impact on infant mortality are mother's current working status, mother's age at the time of first birth, whether the husband lives with her or not, source of drinking water, place of delivery and work status of father. Infant mortality for the male and female infant are 3.87% and 5.06% respectively. This result significantly indicates that male infants are more vulnerable to infant mortality risk than the female infants. Infant mortality rates are 5.35% and 1.18% for the group of infants that received vitamin A dose in the first 2 months after delivery or not received respectively. This result suggests that vitamin A dose has strongly highly significant effect on infant mortality. The availability of good sanitary facilities for the household members has a significant effects on infant mortality. For

instance, infant mortality are 3.87% and 5.00% for the household where the good sanitary facilities available and not available respectively. The number of antenatal visits during pregnancy and type of person who gave prenatal care to the mother prior to the birth also have significant effect on the risk of infant mortality. Birth order number has also significant effect on infant mortality. It was found that 5.41% infant died among the first birth, 3.82% infant died among the 2nd 3rd birth and 4.65% infant died 4th and subsequent birth. Tetanus injection taken before the pregnancy and after the pregnancy also have highly significant effect on risk of infant mortality. For instance, 8.21% of infant died for the mother who did not receive tetanus injection during pregnancy on the other hand only 1.68% infant died for the mothers who received tetanus injection during their pregnancy. Single or multiple birth has a strongly highly significant effect on risk of infant mortality. In case of multiple birth and single birth the infant mortality rate was found 25% and 4.11% respectively. The person who giving the postnatal care on mothers and infants also significantly affect the risk of infant mortality.

Table 3.2 illustrates the relative risk of infant mortality against some socio-economic and demographic characteristics based on simple logistic regression. The result suggested that the risk of infant mortality decreases as mother's education increases. For instance, risk of infant mortality 20%, 30% and 70% lower among the mothers having primary, secondary and higher education than those having no education respectively. It may be for the reason that maternal education insist them to be aware of health facilities, child care and feeding strategies (Cleland and Van Ginneken, 1988). The similar trend also found for the paternal education also. Infant mortality risk for male infant is significantly higher (27%) as compared to female infant (Chen, Huq and D'Souza, 1981). This phenomenon is common for most of the developing countries and true for all populations for the neonatal infant mortality due to biological factors. The level of infant mortality among the household with access to sanitary latrine is 24% lower than that of the households with no access to sanitary latrine (NI Mondal, K Hossain, K Ali, 2009). Receiving tetanus injection during the pregnancy or before the pregnancy significantly influence the infant mortality risk. For instance, infant mortality level among the mothers who received tetanus injection during the pregnancy is 81% lower than those who did not receive tetanus injection during pregnancy (S. Gunasekara). Multiple or single birth has significant effect on infant mortality. Result indicates that children born multiple birth were more than seven-times as likely to die during infancy as those born singletons (odds = 7.77). Person giving postnatal care on mother and infant also significant effect on risk of infant mortality. The result shows, infant mortality is 57% and 56% lower for mother and infant who receive postnatal care by medical personnel than those received by non-medical personnel respectively.

Table 3.4 represents the results of multiple logistic regression analysis of infant mortality on the basis of 11 covariates. These 11 covariates are chosen by stepwise regression which is based on AIC among all independent variables in table 1.1. This 11 covariates also identified through table 3.3 also. Form table 3.4 it is evident that all the predictors significantly affects the risk of infant mortality except the predictor, person giving postnatal care on kids.

4.2 Conclusion

This study was designed to identify the predictors that significantly influence the risk of infant mortality in Bangladesh. The cross tabulation and simple logistic regression analysis suggest that among the 21 predictors only the 13 predictors, mother's & father's education, sex of infant, vitamin A dose, sanitation, antenatal visits during the pregnancy, type of person giving prenatal & postnatal care to mother, birth order number, tetanus injection during pregnancy & before pregnancy, birth type and person giving postnatal care on kids are significant predictors of infant mortality risk. These also support the logical hypothesis as well.

On the other hand, in the multiple logistic regression analysis only the 11 predictors, father's education, mother's current work status, sex of infant, delivery by cesarean section, vitamin A dose, place of delivery, type of person giving prenatal care to the mother, birth order number, Tetanus injection during pregnancy, birth type and person giving postnatal care on kids are identified as significant predictors of risk of infant mortality. Only the 11 predictors has been included in the model and 10 predictors are identified as significant predictors of infant mortality. This situation may be occurred due to the interrelationship among the predictors.

Therefore, the following recommendation can be suggested on the basis of this study:

Infant mortality will decrease if the education facilities increased extensively. It is obvious that education create consciousness and consequently, parents would make proper use of modern health facilities, both for preventive and curative.

Mothers should be on leave from her work during the infant bringing up period

Medical facilities should be provided to the door steps, so that mothers can visit doctors, receive tetanus injections and kids can easily receive vitamin A dose and

postnatal care by the medical personnel.

Appendix A

Bibliography

World Population Prospects: The 2015 Revision. Population Division, United Nations, New York, U.S.A.

Bangladesh Demographic and Health Survey 2011 (BDHS-2011). National Institute of Population Research and Training (NIPORT), Mitre and Associates, Dhaka, Bangladesh and The DHS Program, ICF International, Rockville, Maryland, U.S.A.

McCullagh, P. Nelder, J.A. (1989). Generalized Linear Models, 2nd Ed. Chapman Hall, London

Rawlings, J.O., Pantula, S.G., Dickey, D.A.: Applied Regression Analysis - A Research Tool, 2ed, Springer.

Agresti, A. An Introduction To Categorical Data Analysis, 2ed Wiley, 2007.

Ronald Christensen : Log-Linear Models and Logistic Regression (Springer Texts in Statistics), 2nd ed. 1997.

Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika. 1991;78(3):691–692.

Cleland, J. G. and J. K. Van Ginneken, 1988. Maternal Education and Child Survival in developing Countries: The Search for Pathway of Influence. Soc. Sci. Med., 27: 1357-1368.

Chen, L. C., E. Huq and D'Souza, 1981. Sex Bias in the Family Allocation of Food and Health Care in Rural Bangladesh. Popul. and Develop. Review 7, No. 1 : 55-70.

NI Mondal, K Hossain, K Ali - Journal of Human Ecology, 2009. Factors influencing infant and child mortality: A case study of Rajshahi District, Bangladesh. J Hum Ecol, 26(1): 31-39

S. Gunasekaran: Determinants of Infant and Child Mortality in Rural India, Kalapaz Publication, 2008, p 265.