ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAMME

DEPARTMENT OF STATISTICS

NEONATAL MORTALITY IN ETHIOPIA

GAROMA WAKJIRA

A THESIS SUBMITTED TO THE OFFICE OF GRADUATE STUDIES OF ADDIS ABABA UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN STATISTICS

ADDIS ABABA UNIVERSITY

ADDIS ABABA, ETHIOPIA

JUNE, 2014

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared by Garoma Wakjira, entitled: Neonatal mortality in Ethiopia and submitted in Partial fulfillment of the requirements for the Degree of Master of Science in Statistics complies with requirements of the university and meets the accepted standards with respect to originality and quality.

Signed by the Examining committee:

Examiner______________________Signature__________________Date____________

Examiner ______________________Signature_________________ Date____________

Advisor_______________________ Signature _________________Date____________

____

Chair of Department or Graduate Program Coordinator

ABSTRACT

Neonatal Mortality in Ethiopia

Garoma Wakjira

Addis Ababa University, 2014

Neonatal period, the period during the first 28 days carry one of the highest risk of death in the history of human life. The study aims to identify determinants of neonatal mortality in Ethiopia focusing on socioeconomic, demographic, nutritional and maternal health care seeking predictors. The study used data from the nationally representative 2011 EDHS. Information about 6,303 live-born neonates born within five years preceding the 2011 survey was examined. Both bivariate and multiple logistic regression statistical analyses were used to identify determinants of neonatal mortality in Ethiopia. The chance of neonatal mortality was significantly higher for: neonates born in Amhara (AOR=3.120; 95% CI: 1.288-7.558) and Benishangul-Gumuz (AOR=3.039; 95% CI: 1.242-7.438) regions compared to neonates born in Addis Ababa; neonates belonging to poor households (AOR =1.517; 95% CI: 1.139-2.021); neonates born to mothers in age the group 15-19 (AOR=4.106; 95% CI: 2.412-6.987); 1-4 ranked neonate (AOR=1.751; 95% CI: 1.318-2.326); neonates born within shorter than < 24 months of the preceding birth (AOR=2.666; 95% CI: 2.003-2.549); large size neonates at birth (AOR=1.854; 95% CI: 1.423-2.415); neonates born to mothers who made no ANC visit during pregnancy (AOR=1.624; 95% CI: 0.458-0.851) and neonates delivered at health facilities (AOR=1.790; 95%CI: 1.249-2.564). The odds of neonatal mortality was significantly lower for single birth neonates (AOR=0.075; 95% CI: 0.051-0.109), neonates born in Dire Dawa (AOR=0.888; 95% CI: 0.319-2.469) and neonates born to mothers in the age group 20-34 (AOR=0.841; 95% CI: 0.620-1.141). The study concluded that public health interventions directed at reducing neonatal mortality in Ethiopia should give special attention to predictors which significantly influence neonatal mortality. Short birth interval neonates, child-bearing at early ages and antenatal care utilization should be taken into account when planning the interventions to reduce neonatal mortality in Ethiopia.

ACKNOWLEDGMENT

First and for most, I would like to praise the Lord who has helped me thus far at that dark and troublesome time I passed through and did not leave me alone in that temptation.

I would also like to express my deepest gratitude to my adviser Professor M.K Sharma for his constructive comment and advice throughout the work of this manuscript.

I would also like to appreciate my friends, especially Gutu Adugna who is my former class mate for his encouragement and help throughout the work this thesis. I am also indebted to appreciate my brothers Ayela Wakjira and Kalbessa Gerba for their encouragement and support during my study.

Finally, I would like to extend my heartfelt thanks to Department of Statistics, Addis Ababa University, for all their support during my stay at this department.

LIST OF ACRONYMS

AIC	Akaike's information criterion
AIDS	Acquired Immune Deficiency Syndrome
ANC	Antenatal Care
AOR	Adjusted Odds Ratio
CSA	Central Statistical Agency
DHS	Demographic and Health Survey
EA	Enumeration area
EDHS	Ethiopia Demographic and Health Survey
HIV	Human Immunodeficiency Virus
MDG	Millennium Development Goal
MOH	Ministry of Health
OR	Odds ratio
ROC	Receiver operating characteristics
SC	Schwarz information criterion
SNNP	Southern Nations, Nationalities and Peoples
UNICEF	United Nations Children's Fund
UN-IGME	United Nations Inter-agency Group for Child Mortality Estimation

Dedication

This work is dedicated to the memory of every new born child whose sojourn on earth lasted only the first 28 days of life.

Table of Contents

Contents Page

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 Background

Neonatal period, the period during the first 28 days carry one of the highest risk of death in the history of human life. Whereas developed countries maintain a neonatal mortality rate of 2-3/1,000, it is not uncommon that this rate reaches over 60/1,000 in the poorer segments of the world (Rajaratnam *et al.*, 2010). Most neonatal deaths (99%) arise in low and middle-income countries, and about half occur at home especially in sub-Saharan Africa and South Asia (Malqvist, 2011). More than one-third of the neonatal deaths in the world occur in three south Asian countries - India, Pakistan and Bangladesh. Among these countries, India has the largest number of neonatal deaths primarily because of large number of births. The top ten countries of the world which contribute 67% of neonatal deaths are India (27%), China (10%), Pakistan (7%), Nigeria (6%), Bangladesh (4%), Ethiopia (4%), Democratic Republic of the Congo (3%), Indonesia (2%), Afghanistan (2%) and United Republic of Tanzania (2%) (Lawn *et al*., 2005).

The reduction in neonatal mortality rates has been slower, compared to both under-five and child mortality rates after the first month of life (Titaley *et al*., 2008). The good news is that neonatal mortality is on the decline globally. The world‟s neonatal mortality rate fell from 33 deaths per 1,000 live births in 1990 to 21 per 1,000 in 2012. All regions saw drops, with lower percentage reductions in South Asia and sub-Saharan Africa (39% and 28% respectively) than other regions; even the smaller regional reductions represent significant progress. The overall result was a reduction of global neonatal deaths from 4.6 million in 1990 to 2.9 million in 2012 (UNICEF, 2013).

Millennium Development Goal 4 (MDG 4) calls for reducing the under-five mortality rate by two thirds between 1990 and 2015. As global momentum and investment for accelerating child survival grow, monitoring progress at the global and country levels has become even more critical. The most recent Inter-agency Group for Child Mortality Estimation (IGME) estimates show that nearly 8.1 million under-five children died in 2009, implying that more than 22,000 children die per day. These figures reflect substantial progress. Globally, the under-five mortality rate has fallen from 89 deaths per 1,000 live births in 1990 to 60 deaths per 1000 live births in 2009. But the rate of decline –a one-third reduction over 20 years –is insufficient to meet MDG 4, particularly in Sub-Saharan Africa, Southern Asia and Oceania. Neonatal and post-neonatal mortality rates declined less, 3.0% and 2.5% decrease respectively. As a consequence there is an increasing proportion of infant deaths occurring in the neonatal period worldwide, which now accounts for two-thirds of deaths in children less than one year old, and nearly four-tenths of all deaths in children less than five years of age (Moss *et al.*, 2002).

Infant and child mortality rates reflect a country‟s level of socioeconomic development and quality of life. They are used for monitoring and evaluating population and health programmes and policies. The rates are also important for monitoring progress towards the United Nations MDG to reduce child mortality as expected by the year 2015. It would be difficult to achieve MDG without reducing child mortality by two-thirds by 2015 (Haines et al, 2004). Thus the study on infant and child mortality; in particular, neonatal mortality is of great importance to monitoring the progress of children‟s health status.

With a population of nearly 83 million in 2010, Ethiopia is the second most populous country in Africa after Nigeria. The population grows at a rate of 2.6% per annum and the majority of people (84%) reside in rural areas, with agriculture being the major source of livelihood. High mortality, high fertility and low life expectancy characterize the demography, as in most sub-Saharan Africa countries. In the past decade, however, the country witnessed an unprecedented decline in under-5 mortality from 166 per 1000 in 2000 to 88 per 1000 live births in 2011, an average decline of 47 % (CSA & ICF international, 2012).

In Ethiopia, results from the 2011 EDHS data showed a remarkable decline in all levels of childhood mortality. The same report showed that infant mortality has declined by 42 percent over the 15-year period preceding the survey from 101 deaths per 1,000 live births to 59 deaths per 1,000 live births. Furthermore, under-five mortality has declined by 47 percent over the same period from 166 deaths per 1,000 live births to 88 deaths per 1,000 live births. Even though not to the same extent, the neonatal mortality has also decreased over the 15-year period preceding the survey by 31 percent from 54 deaths per 1,000 live births to 37 deaths per 1,000 live births. This reduction in neonatal mortality, as in other parts of the world, was slower than for infant, and under-five mortality, which fell by 42 percent and 47 percent respectively over the 15 year period (CSA & ICF international, 2012). In addition, the country is experiencing a high neonatal mortality rate at 37 per 1000 live births, comparable to the average rate of 35.9 per 1000 live births for the African region overall (Oestergaard *et al.*, 2011).

In Ethiopia neonatal mortality has declined by 21% and mortality under the age of five declined by 26% between 2000 and 2005 (Susman, 2011). According to the results of the 2005 Ethiopia Demographic and Health Survey (EDHS), infant mortality in urban areas was found to be 66 deaths per 1,000 live births compared to 81 deaths per 1,000 live births in rural areas giving an average of 77 infant deaths per 1,000 live births for the country. Approximately half of infant deaths in Ethiopia occur during the first month of life (CSA and ORC Macro, 2006).

Mortality trends can also be examined by comparing data from DHS surveys conducted in 2000, 2005, and 2011. Infant and under-five mortality rates obtained by these surveys evidence a continuous declining trend in mortality. Under-five mortality decreased from 166 deaths per 1,000 live births in the 2000 survey to 88 in 2011, while infant mortality decreased from 97 deaths per 1,000 live births in the 2000 survey to 59 in the 2011 survey. On the other hand, even though neonatal mortality rate decreased from 49 deaths per 1,000 live births in 2000 to 39 deaths per 1,000 live births in 2005, it has since remained stable at 37 deaths per 1,000 (CSA & ICF international, 2012).

Evidences suggest that while neonatal health is found to be dependent on health care services, post-neonatal health is dependent largely on environmental factors. It is believed that the factors associated with infant mortality are many. Their relative importance varies across populations depending upon the level of socio-economic and environmental set up.

Some of the most important factors associated with infant mortality are premature and low birth weight, health problems particularly respiratory infections; socio-economic and demographic factors such as class, mother‟s age, birth weight and interval from previous delivery; and maternal and child health care variables such as breastfeeding, nutritional status, place of delivery and type of delivery care (Regassa, 2012).

There is a general dearth of studies on neonatal mortality in Ethiopia, which significantly limits our understanding of the breadth and depth of the problem for evidence-based programming. The causes of neonatal mortality are not well documented in Ethiopia, but previous studies report causes such as sepsis, asphyxia, birth injury, tetanus, preterm birth, congenital malformations and unknown causes (World Bank, 2013). Hence, this study attempts to identify the determinants of neonatal mortality in Ethiopia focusing on socio-economic, demographic, nutritional and maternal health care seeking factors.

1.2 Statement of problem

Millennium Development Goal 4 (MDG 4) calls for reducing the under-five mortality rate by two-thirds between 1990 and 2015. The world has made substantial progress, reducing the under-five mortality rate by 47 percent, from 90 deaths per 1,000 live births in 1990 to 48 in 2012. However, this progress has not been enough, and the target is at risks of being missed at the global level. To achieve MDG 4 on time, the global annual rate of reduction in under-five mortality rate would need to rise to 15.6 percent for 2012–2015, much faster than the 3.9 percent achieved over 2005–2012 (UN IGME, 2013).

Over one third of the global 10.8 million deaths of children under age of five in 2000 occurred in the neonatal period (Black et al., 2003). Declines in neonatal mortality over the last three decades have been slower than declines in post-neonatal or early child mortality. In order to achieve the Millennium Development Goal target of a two thirds reduction in under-five mortality from 1990 to 2015 (UN, 2001), major reductions are going to be required in neonatal mortality (Lawn et al., 2005). However, few developing countries have vital registration systems that are complete enough to provide accurate estimates of neonatal mortality. Causes of death in the neonatal period in the developing world are poorly measured also, though major components are believed to be birth asphyxia, severe infections, complications of prematurity and tetanus (Lawn *et al*., 2005).

Over the last decade, neonatal mortality has gained importance on the world policy agenda because the Millennium Development Goal (MDG) for child survival cannot be met without substantial reductions in neonatal mortality. It is estimated that reduction of under-5 child mortality by two-thirds by 2015, as called for by the MDG, requires a reduction in neonatal mortality of at least 50% (Hyder *et al*., 2003). There are highly feasible and cost-effective interventions that could avert up to 72% of neonatal deaths, but this can only be achieved if countries adopt locally relevant and focused interventions that are guided by evidence.

Many studies have shown that neonatal mortality is influenced by multiple factors. Maternal health before, during, and after pregnancy, conditions at the time of labor and delivery and postnatal care of babies play a significant role in reducing neonatal mortality. Socioeconomics, demographics, the health care system, cultural practices and technology are also important indirect determinants of neonatal mortality (Lawn *et al*., 2005, Rahman *et al*., 2012, Rahman *et al*., 2010). Thus, this study attempts to answer the question that what are the socioeconomic, demographic, nutritional and maternal health care services determinants of neonatal mortality in Ethiopia.

1.3 Objectives

1.3.1 General Objective

The general objective of this study is to identify socio-economic, demographic, nutritional and maternal health care seeking determinants of neonatal mortality in Ethiopia.

1.3.2 Specific objectives

- ✓ Examine the association between independent variables (covariates) and neonatal mortality.
- ✓ Make relevant recommendations for concerned bodies based on the findings of the study regarding the determinants of neonatal mortality in Ethiopia.

1.4 Significance of the study

The results of this study may provide information on determinants of neonatal mortality in Ethiopia by analyzing the effect of different covariates on neonatal mortality.

Specifically

- The results of this paper are expected to give some knowledge about determinants and risk factors of neonatal mortality in Ethiopia.
- The results of this paper could be used as input for other studies related to neonatal mortality.
- The findings of this study may help policy makers, programme managers and donors in driving up understanding about the determinants of early childhood mortality, particularly neonatal mortality in Ethiopia.

1.5 Organization of the Thesis

The thesis comprises five chapters. The first chapter (meaning this one) introduces and provides a description of the background, statement of problem, objectives, significance, organization and limitation of the study. Chapter two presents literature review related to neonatal mortality. Chapter three presents the data and methodology used in the study. The fourth chapter presents the results and discussion of the study. Chapter five is about the conclusions and recommendations of the study.

1.6 Limitation of the study

- The information used in this study is based on the retrospective birth history of children and reported characteristics of mothers and households by the mother‟s of the children, which may cause under-reporting of information.
- Only surviving women were interviewed, which may have led to an underestimate of the neonatal mortality rate, because of the association of neonatal deaths with maternal deaths. This could also have led to an underestimation of the effect of some of the associated factors, such as delivery complications.
- There are other possible determinants of neonatal mortality, which were not available in the EDHS dataset, such as gestational age, essential newborn care practices.

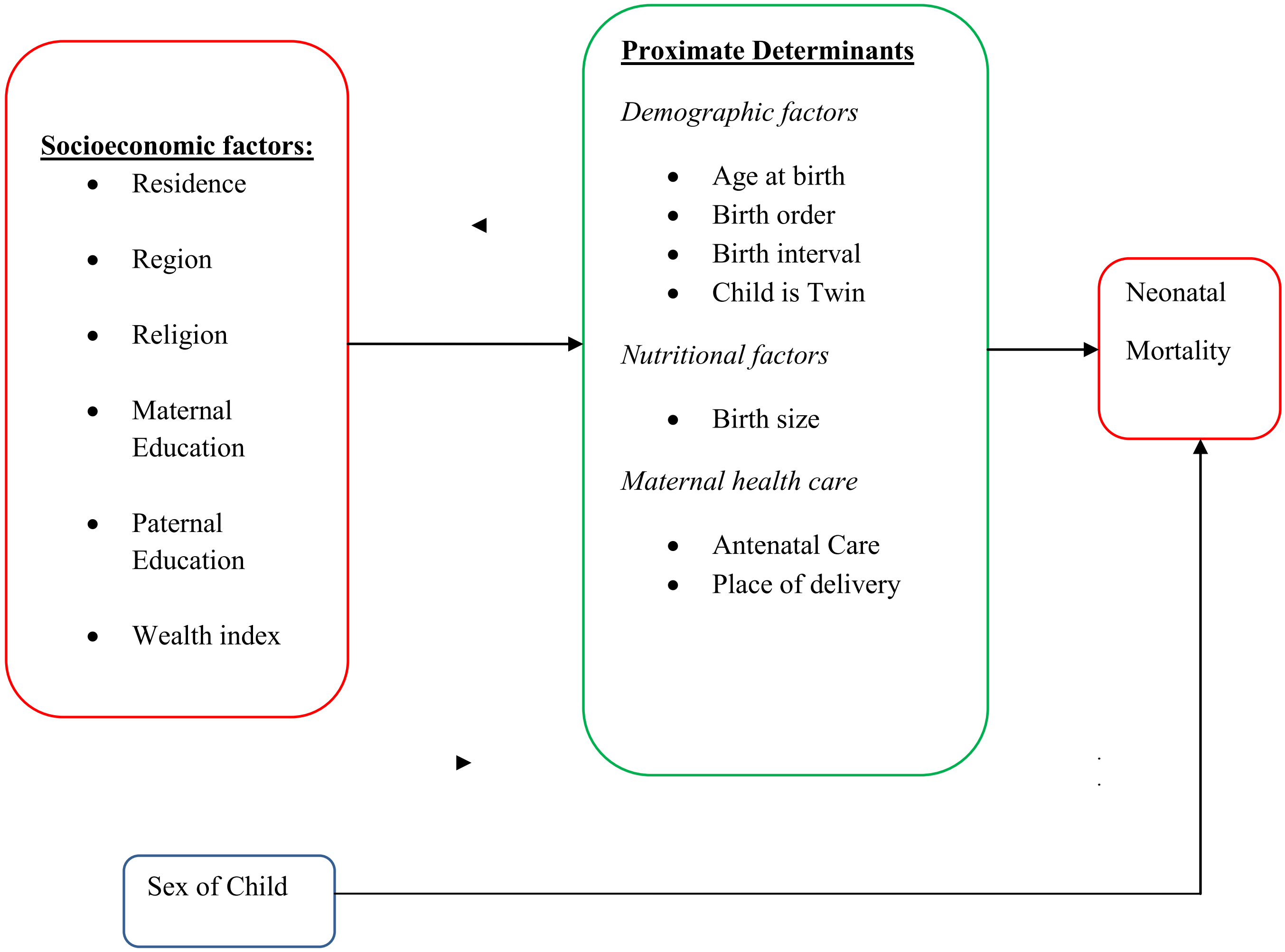
CHAPTER TWO

LITERATURE REVIEW

2.1 Conceptual framework

Many researchers such as Mahmood (2002) and later on Titaley *et al*. (2008) adapted the conceptual framework proposed by Mosley and Chen (1984) with some modifications based on the limitations and structure of the DHS data. Keeping in view the aforementioned frameworks, this study considered neonatal mortality as the outcome variable. The socioeconomic factors affect the outcome variable through proximate determinants namely; demographic factors, nutritional factors and maternal health care seeking factors.

The conceptual framework in the figure below highlights the interrelationships between the variables considered in this study. It will be observed that in line with Mosley and Chen‟s classical framework, socioeconomic factors are located most distal to neonatal mortality, i.e., its influence on neonatal mortality is through proximate factors (indirect). These factors are connected through a series of pathways (arrows) to other more proximal factors (demographic, nutritional and maternal health care seeking factors) which in turn are linked to neonatal mortality. Socio-economic factors rarely cause ill health or mortality directly hence most of their influence on neonatal mortality is largely exerted through associations with more proximal factors. Note that sex of the child is placed independently in the framework since it is not determined by other socioeconomic factors but may be associated either directly or indirectly with neonatal mortality.



Conceptual framework for factors influencing neonatal mortality, adapted from Mosley and Chen (1984).

2.2 Empirical literature

Empirically, many studies have shown that neonatal mortality is influenced by a number of socio-economic, demographic, nutritional and maternal health care seeking factors.

For instance, Kamal (2012) used data from Bangladesh Demographic and Health Survey (BDHS) conducted in 2007 to investigate the effect of maternal education on neonatal mortality in Bangladesh. The study used both bivariate and multivariate statistical analyses to assess the relationship between neonatal mortality and contextual factors focusing on maternal education. The prevalence of neonatal mortality was 37/1,000. The sequential multivariate logistic regression analyses yielded strong significant negative association between maternal education and neonatal mortality. The result also revealed that maternal age, religion, birth order and antenatal care seeking are also important determinants of neonatal death.

Chowdhury *et al*. (2013) conducted another study in Bangladesh using the data extracted from the 2007 BDHS. Multivariate proportional hazard models are employed to study determinants of neonatal and post-neonatal mortality. The results show that father‟s education, place of residence, housing materials, number of children under five years of age, and previous death of sibling have significant influence on neonatal mortality. The findings also indicate that mother‟s education, toilet facility, number of children under five and breastfeeding have significant effect on post-neonatal mortality in Bangladesh.

Lawn *et al*. (2005) observed that every year an estimated 4 million babies died in the first 4 weeks of life. A similar number are stillborn and half a million mothers died from pregnancy-related causes. The results of the study indicate that the main direct causes of neonatal death are estimated to be preterm birth (28%), severe infections (26%), asphyxia (23%) and neonatal tetanus (7%). Low birth weight is an important indirect cause of death. Maternal complications in labour carry a high risk of neonatal death. They concluded that poverty is strongly associated with an increased risk.

Patel and Sharma (2013) conducted a study using data from Indian National Family Health Survey -3 (NFHS-3) to find the association of neonatal mortality rate with different risk factors and association of neonatal mortality rate with various interventional variables. The sample for analysis includes all 29 states of India in which the third round of National Family Health Survey were conducted and reports for the same were available. Neonatal mortality rates of different states were taken as dependent variable. The result of bivariate regression showed the influence of percentage of women age 15-19 years who have begun childbearing and percentage of women with BMI < 17 on neonatal mortality was confounding effect of socio-economic status. The bivariate regression also showed the influence of antenatal check up, iron-folic acid supplementation and postnatal check up on neonatal mortality was confounding effect of socio-economic status.

Chaman et al. (2009) conducted a study to evaluate neonatal mortality risk factors based on nested case-control design. The study population was 6900 neonates who were born in rural areas of Kohgiluyeh and Boyerahmad province, south Iran. 97 cases and 97 controls were selected in study cohort. The results revealed that Prematurity, low birth weight, caesarean section, birth rank more than three and birth spacing less than 24 months showed significant statistical association ($P < 0.05$) with neonatal mortality.

Kousar *et al*. (2010) carried out a study at the department of pediatrics/nursery, Liaquat University Hospital Hyderabad, India, from July 2006 to December 2007 to identify the common causes and risk factors of neonatal mortality in a tertiary care Hospital. A total of 1,203 neonates were admitted in the nursery during the study period; out of them 284 expired. These deaths were evaluated for age, gender, gestational age, birth weight, risk factors and cause of death. The result showed that out of 284 neonates 218 were low birth weight and 66 full term normal. Most deaths occurred to *males* except in preterm SGA females. The main risk factors identified were maternal anemia, maternal malnutrition, poor antenatal care, grand multi-parity, Premature Rapture of membranes (PROM) followed by maternal fever Ante partum Hemorrhage (APH) and maternal Urinary Track Infection (UTI). The most common causes of death were sepsis, birth asphyxia, respiratory distress syndrome and congenital anomalies neonates.

Andersson *et al.* (2002) conducted a study in a rural community in Ethiopia that has been under monthly demographic surveillance since 1987 based on data collected in the first decade (1987-97) to calculate mortality incidence rates and analyze survival. The result showed that the overall mortality rate was 27/1000 live births. The result also showed that the rates in the early and late periods were 20 and 8/1000 live births, respectively. The mortality incidence rates show that, every day, three of every 1000 newborns died in the first week of life. Neonatal mortality accounted for 43% of infant mortality. If all neonates survived the 1st week of life, life expectancy would increase by one year. The increased risk of neonatal mortality was found to be associated with living in a *rural lowland area, twin birth and male gender*.

Most neonatal deaths result from preterm birth or delivery complications. Home delivery without a skilled health care provider leaves women and infants at greater risk of these complications. Improving quality and access to primary health care throughout pregnancy and birth is thus essential. High-impact interventions during delivery and the postnatal period, such as safe and clean delivery, skilled attendance at birth and exclusive breastfeeding, can drastically reduce neonatal morbidity and mortality (Zuma *et al*., 2013).
The study aimed to identify the determinants of neonatal mortality in Indonesia, for a nationally representative sample of births from 1997 to 2002 was conducted by Titaley *et al* (2008). The data source for the analysis was the 2002–2003 Indonesia Demographic and Health Survey from which survival information of 15,952 singleton live-born infants born between 1997 and 2002 was examined. Multilevel logistic regression using a hierarchical approach was performed to analyze the factors associated with neonatal deaths, using community, socio-economic status and proximate determinants. The results revealed that at the community level, the odds of neonatal death was significantly higher for infants from East Java, and for North, Central and Southeast Sulawesi and Gorontalo combined compared to the lowest neonatal mortality regions of Bali, South Sulawesi and Jambi provinces. A progressive reduction in the odds was found as the percentage of *deliveries assisted* by trained delivery attendants in the cluster increased. The odds of neonatal death were higher for infants born to mothers and fathers who were both employed, and for infants born to fathers who were unemployed. The odds were also higher for higher rank infants with a short birth interval, male infant, smaller than average size infants, and infants whose mother had a history of delivery complications. Infants receiving any postnatal care were significantly protected from neonatal death.

Singh *et al*. (2013) conducted a study to examine individual, household, and community level factors affecting neonatal mortality in rural India. The information on 171,529 singleton live-births from the most recent round of the District Level Household Survey conducted in 2007–08 was analyzed. Two-level logistic regression was performed to analyze factors associated with neonatal deaths in rural India. The results showed that the odds of neonatal death were lower for neonates born to mothers with secondary level education compared to those born to illiterate mothers. A progressive reduction in the odds occurred as the level of fathers‟ education increased. The odds of neonatal death were lower for infants born to unemployed mothers compared to those who worked as agricultural worker/farmer/laborer. The odds decreased if neonates belonged to Scheduled Tribes or „Others‟ caste groups and to households with access to improved sanitation, pucca house and electricity. The odds were higher for male infants and whose mothers experienced delivery complications. Infants whose mothers received two tetanus toxoid injections were less likely to die in the neonatal period. Children of higher birth order were less likely to die compared to first birth order.

Diallo *et al*. (2011) conducted a study to measure the neonatal mortality rate (NMR) and investigate its predictors in a rural area of Burkina Faso. A cohort of infants born in 24 villages in Banfora region was followed until the children were 6 months old. Among 864 live births followed to day 28, there were 40 neonatal deaths, a NMR of 46.3 per 1000 live births. Multivariable regression identified twin birth, having a nulliparous mother and birth into a polygynous household as the main predictors of neonatal death.

Kamal (2011) conducted a study which aims to investigate the effect of adolescent motherhood on neonatal mortality in Bangladesh using data from the nationally representative 2007 BDHS. Both univariate and multivariate statistical analyses were used to assess the relationship between neonatal mortality and socio-demographic contextual factors focusing on maternal age in particular, adolescent and adult motherhood. The sequential multivariate logistic regression analyses yielded significantly increased risk of neonatal mortality among children of adolescent mothers than of adult mothers. Maternal education, religion, birth rank, and antenatal care seeking are also important determinants of neonatal death.

Zwane and Masango (2012) conducted a study *with the aim of examining* the effects of socioeconomic and maternal variables on the probability of neonatal deaths in Swaziland. The data source was the 2006-07 Swaziland Demographic and Health Survey from which survival information on 1,727 infants born within the 3 years preceding the survey. Design based logistic regression incorporating survey weights was performed to analyze associated factors. Compared to infants born at home, the odds of dying were significantly lower for infants born in a private facility. Neonates born in public facilities and those born at home had similar odds of dying. For newborns, whose birth size according to the mother was smaller than average, the odds of dying were more than 4 times the odds for large-sized babies.

Susuman (2012) analyzed data from the Ethiopia Demographic and Health Surveys 2000 and 2005 using indirect estimation of Brass and Trussell. The results showed that neonatal and post-neonatal mortality declined gradually. *Birth intervals shorter than 2 years* led to higher neonatal mortality rates than higher birth intervals. The study concluded that proper spacing of births would provide more time for childcare by making more maternal resources available for the care of the child and mother.

Tesfaye (2003) conducted a community based cross sectional comparative study in Dire-Dawa, Ethiopia, during November 2002 through April 2003 to assess the pregnancy outcome with emphasis on perinatal and neonatal mortality by delivery place and its associated factors in Dire Dawa town. A total of 1,462 mothers who had children or had been pregnant for the last five years participated in the study. Pretested standardized questionnaires were used to obtain information on socio demographic, obstetric history and the condition of the new born and mothers during labor and neonatal period. The *findings* of the study revealed that mothers who had 2-4 parity, had higher risk of perinatal mortality than primipara mothers and mothers who had 5+ parity had higher risk of perinatal mortality than primipara mothers. The study found that very small and small (mother‟s perceptions birth weight) neonates had higher risk of neonatal mortality than neonates who had normal birth weight at birth and also term babies had lower risk of perinatal and neonatal mortality than babies born preterm. Mothers‟‟ income also showed significant impact on neonatal mortality in this study.

Seedhom and Kamal (2010) conducted a community based study to determine neonatal mortality rate (NMR), risk factors and its causes in a rural area, El- Minia governorate, Egypt. Only mothers who gave birth in the year 2008 were included. A questionnaire had been designed inquiring about possible risk factors of neonatal death and medical care received by infants and their mothers. The result revealed that the NMR was 24/1000 live births. Maternal demographic characteristics were associated significantly with neonatal mortality as 27.1%, 83.3%, and 79.2% of neonatal deaths occurred with maternal age < 20 years, maternal illiteracy, and no breast feeding respectively. Main causes of death were low birth weight, prematurity, and infection.

Araújo et al. (2000) conducted a cohort study to establish the profile of neonates in Caxias do Sul city, Brazil and to study early neonatal mortality, its causes and related variables. The study enrolled 5,545 newborns, which were followed up to 7 days after birth. The probability of early neonatal mortality was calculated and multiple logistic regression was performed to relate all studied variables to the outcome of early neonatal death. The *results* revealed that the observed probability of early neonatal mortality was 7.44 per thousand live births. The incidence of premature births and low birth weight was 9.4% and 8.1%, respectively. Fifty five percent of the neonates were born through cesarean section. The causes of death were related to socioeconomic and educational level. Previous history of neonatal mortality, maternal age > 35 years, gestational age, Apgar score < 7, male sex and low birth weight were related to early neonatal death. The main cause of death was hyaline membrane disease, followed by congenital cardiopaties, extreme preterm and abruptio placentae.

Chowdhury *et al*. (2010) conducted a study to determine socio-economic factors that affect infancy and childhood mortality**.** The result showed that neonatal mortality rates (NMR), post-neonatal mortality rates (PNMR) and infant mortality rates (IMR) were higher among *illiterate* reproductive mothers and of whom houses have unhygienic latrine. The Chi-square (χ^2) test result showed that mother‟s education; types of latrine and electricity have significant association with neonatal, post neonatal, infant and child mortality. Multivariate analysis results identified *mother's education* and *occupation* as having momentous influence on mortality of post-neonatal period but in infant and child period, *parents' education* and *occupation*, types of latrine and electricity have significant effects on mortality.

The study by Rutstein (2005) examined the association between birth intervals and infant and child mortality and nutritional status using repeated analysis of retrospective survey data from the Demographic and Health Surveys (DHS) program from 17 developing countries collected between 1990 and 1997. The key independent variable was the length of the preceding birth interval measured as the number of months between the birth of the child under study (index child) and the immediately preceding birth to the mother, if any. Both bivariate and multivariate designs were employed. Several child and mother-specific variables were used in the multivariate analysis in order to control for potential bias from confounding factors. Adjusted odds ratios were calculated to estimate relative risk. The results revealed that for neonatal mortality and infant mortality, the risk of dying decreases with increasing birth interval lengths up to 36 months, at which point the risk plateaus. For child mortality, the analysis indicates that the longer the birth interval, the lower the risk, even for intervals of 48 months or more. The relationship between chronic malnutrition and birth spacing is statistically significant in 6 of the 14 surveys with anthropometric data, and in 5 surveys between general malnutrition and birth spacing. However, there is a clear pattern of increasing chronic and general undernutrition as the birth interval is shorter, as indicated by the averages of the adjusted odds ratios for all 14 countries.

CHAPTER THREE

DATA & METHODOLOGY

3.1 Source of Data

The data for this study have been taken from the 2011 Ethiopia Demographic and Health Survey (EDHS) which is the third DHS in Ethiopia, following the 2000 and 2005 EDHS. The 2011 EDHS was carried out under the aegis of the Ministry of Health (MOH) and was implemented by the Central Statistical Agency (CSA). The principal objective of the 2011 EDHS is to provide current and reliable data on fertility and family planning behavior, child mortality, adult and maternal mortality, children‟s nutritional status, use of maternal and child health services, knowledge of HIV/AIDS, and prevalence of HIV/AIDS and anemia.

The sample for the 2011 EDHS was designed to provide population and health indicators at the national and regional levels. The sampling frame used for the 2011 EDHS was the Population and Housing Census conducted by the Central Statistical Agency (CSA) in 2007. The 2011 EDHS sample was selected using a stratified, two-stage cluster design and Enumeration areas (EAs) were the sampling units for the first stage sampling. The 2011 EDHS sample included 624 EAs, 187 in urban areas and 437 in rural areas.

Households comprised the second stage of sampling. A complete listing of households was carried out in each of the 624 selected EAs from September 2010 through January 2011. A representative sample of 17,817 households was selected for the 2011 EDHS. Of these, 16,702 were successfully interviewed, yielding a household response rate of 98 percent. In the interviewed households, 17,385 eligible women (15-49) were identified for individual interview; complete interviews were conducted for 16,515, yielding a response rate of 95 percent. Similarly, a total of 15,908 eligible men (15-59) were identified for interview; complete interviews were conducted for 14,110, yielding a response rate of 89 percent.

3.2. Study Variables

Response (outcome) variable: the outcome variable of this study is „neonatal death‟, which is defined as the death of a live-born infant within 28 days of life which can be recorded as binary (1= died, 0= not died).

Explanatory variables

Table 3.1 lists all explanatory variables, their definitions and categories used in this study. These variables can be divided into socioeconomic (distal) factors and proximate determinants. These variables were selected based on review literatures on the determinants or risk factors of neonatal mortality.

Table 3.1: Operational definition and categorization of the variables used in the analysis

variables	Definition and categorization
Socioeconomic (distal) factors	
Residence	Place of residence: (1) Urban (2) Rural
Region	Administrative regions: (1) Tigray (2) Affar (3)Amhara (4) Oromiya (5) Somali (6) Benishangul-Gumuz; (7) SNNP (8)Gambela (9)Harari(10) DireDawa (11)Addis Ababa
Maternal education	Education: (1) No education (2) Primary (3) Secondary & higher
Paternal education	Education: (1)No education(2) Primary (3) Secondary & higher
Wealth index	Household weald index: (1) Poor(2) medium(3) Rich
Religion	mother‟s religion: (1) Christian (2)Muslim (3)Traditional/Others
Proximate Determinants	
Maternal age	Mother‟s age at child birth: (1) 15-19 (2) 20-34 (3) 35-49
Birth order	Birth rank of child: (1) 1-4 (2) 5 or more
Birth interval	Preceding birth interval(in months): (1) <24 (2) 24-36 (3) > 36
Sex	Sex of the neonate: (1) male (2) female
Type of birth	Type of neonate birth: (1) single (2) multiple
Birth size	Mother‟s perception of birth size at birth: (1) average (2) small (3) Large

ANC seeking	No. of antenatal visits during pregnancy: (1) None (2) < 4 (3) ≥ 4
Place of delivery(POD)	Place of delivery: (1) Home (2) Health facility

3.3 Methodology

Descriptive, bivariate and multiple logistic regression analyses have been used in this study. Bivariate analysis has been applied to examine the association of various independent variables and neonatal mortality. This was followed by multiple logistic regression analysis. Statistical packages SPSS, STATA and SAS have been used for data analysis.

3.3.1 Binary Logistic Regression

Logistic regression is part of a family of models called the Generalized Linear Model used when the response variable is qualitative or categorical in nature and independent variables can be continuous and/ or categorical. Binomial or binary logistic regression is the form of regression which is used when the dependent variable is dichotomous and the independent variables are of any type (Hosmer and Lemeshow, 2000).

Binary logistic regression techniques resolve inconsistencies associated with dichotomous dependent data and the assumptions of linear regression methods. The independent variables that are used for outcome prediction may be dichotomous, categorical or continuous. Binary logistic regression is commonly used in manufacturing and health related studies. It can be used for any application where binary outcomes can be predicted. Logistic regression is based on the logit transformation of the dependent variable. The logit transformation generates a continuous logarithmic curve from non-continuous data so that a regression model can be developed. The outcome probabilities for each dependent variable value are the basis for the model. The logit transformation is necessary since dichotomous dependent data violates ordinary least squares assumptions. Another issue with dichotomous data is that the error terms are not normally distributed, thus ordinary sum of squares regression and all normality tests are invalid (Healy, 2006).

Logistic regression is less restrictive than linear regression. It does not require normally distributed dependent data or homogeneity of variance.

Predictions made by linear regression are based on the observed changes in the independent data itself. Logistic regression is based on the log of the odds of a particular event occurring with a given set of observations. Logistic regression"s underlying principles are based on probabilities and the nature of the log curve.

Discriminant analysis and logistic regression will produce similar results with dichotomous dependent data except discriminant analysis is more restrictive and complex. Unlike discriminant analysis, logistic regression does not restrict the nature of the independent variable. In contrast with discriminant analysis, logistic regression doesn"t restrict categorical independent variables. Discriminant analysis relies on strict adherence to normality and the equal variance assumptions while logistic regression does not have this requirement (Healy, 2006).

Logistic regression has two main uses:

- The first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of the analysis are in the form of an odds ratio.
- Logistic regression also provides knowledge of the relationships and strengths among the variables.

There are two primary reasons for choosing the logistic distribution function. First, from a mathematical point of view, it is an extremely flexible and easily used function, and second, it lends itself to a clinically meaningful interpretation (Hosmer and Lemeshow, 2000).

3.3.1.1 Assumptions

- Meaningful coding. Logistic coefficients will be difficult to interpret if not coded meaningfully. The convention for binary logistic regression is to code the dependent class of greatest interest as 1 and the other class as 0, and to code its expected correlates also as +1 to assure positive correlation.
- The explanatory variables are not linear combinations of each other.
- Logistic regression does not assume a linear relationship between the dependent and independent variables.

- The categories (groups) must be mutually exclusive and exhaustive; a case can only be in one group and every case must be a member of one of the groups.
- The dependent variable need not be normally distributed.
- Logistic regression does not require that the independents be interval.
- Normally distributed error terms are not assumed.
- Larger samples are needed than for linear regression because maximum likelihood coefficients are large sample estimates. A minimum of 50 cases per predictor is recommended (Pampel, 2000, Midi et al., 2010).

3.3.1.2 Odds ratio

Logistic regressions work with odds and odds ratio. The odds are simply the ratio of the probabilities for the two possible outcomes. If π is the probability that the event will occur, then $1-\pi$ is the probability that the event will not occur:

$$Odds = \frac{\pi}{1-\pi} \tag{3.1}$$

In the 2 × 2 contingency table, within row 1 the *odds* of success are $odds_1 = \pi_1/(1-\pi_1)$, and within row 2 the odds of success equal $odds_2 = \pi_2/(1-\pi_2)$. The ratio of the odds from the two rows,

$$\theta = \frac{odds_1}{odds_2} = \frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)} \tag{3.2}$$

is the odds ratio.

3.3.1.3 Logistic Regression Model

When the response variable is binary, there is considerable empirical evidence that the shape of the response function should be nonlinear. A monotonically increasing or decreasing S-shaped or reverse S-shaped function. For a binary response variable Y and an explanatory variable X, let π=P(Y=1/X=x) =1-P(Y=0/X=x). One possible logistic regression model is given by

$\pi = \dfrac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}$, where β_0=intercept and β_1=slope

Thus, if the event of interest occurs, in our case neonatal mortality, with probability

$\pi = \dfrac{e^{\beta_0+\beta_1 \mathrm{x}}}{1+e^{\beta_0+\beta_1 \mathrm{x}}}$, then the odds in favor of success is

$$\frac{\pi}{1-\pi} = \frac{e^{\beta_0+\beta_1 \mathrm{x}}/1+e^{\beta_0+\beta_1 \mathrm{x}}}{1/1+e^{\beta_0+\beta_1 \mathrm{x}}} = e^{\beta_0+\beta_1 \mathrm{x}} \tag{3.3}$$

Taking the natural logarithm of each side of equation (3.3),

$$\mathrm{logit}(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) = \ln\left[e^{\beta_0+\beta_1 \mathrm{x}}\right] = \beta_0 + \beta_1 \mathrm{x} \tag{3.4}$$

Thus, modeling the probability π with logistic function is equivalent to fitting a linear regression model in which the continuous response y has been replaced by the logarithm of the odds of success for a dichotomous random variable. Instead of assuming linear relationship between π and x, we assume the linear relationship between $\ln\left(\frac{\pi}{1-\pi}\right)$ and x. The technique of fitting a model of this form is known as logistic regression.

3.3.1.4 Multiple logistic regression

Logistic regression can be easily extended to the situations with multiple predictors (explanatory) variables. Consider a collection of p independent variables denoted by the vector $\mathbf{x} = (x_1, x_2, ..., x_p)'$. Let the conditional probability that the outcome is present be denoted by P(Y=1| $\mathbf{x}$) $=\pi$. The logit function for multiple logistic regressions is given as follows

$$\mathrm{logit}[\pi] = \ln\left(\frac{\pi}{1-\pi}\right) = \ln(e^{g(\mathbf{x})}) = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p \tag{3.5}$$

in which case

$$\pi = \frac{e^{g(\mathrm{x})}}{1+e^{g(\mathrm{x})}} = \frac{e^{\beta_0+\beta_1 x_1+..+\beta_p x_p}}{1+e^{\beta_0+\beta_1 x_1+..+\beta_p x_p}} \tag{3.6}$$

where a model parameters β_i will be interpreted as the change in the log-odds of a success event (Y=1) for a one unit increase in x_i, holding all the other predictors constant.

In order to estimate the parameters, maximum likelihood-based iteration algorithms can be employed. Inspite of the attractive properties of the logit function, it is by no means the only suitable function for transforming probabilities to arbitrary real values. The general term for such a transformation function is the *link function*, as it links the probabilities (or more generally, the expected values of the dependent variable) to the explanatory variables.

This implies that the logit of the probability of an event given X is a simple linear function. That is, logistic regression is really just a standard linear regression model, once we transform the dichotomous outcome by the logit transform. This transform changes the range of π from 0 to 1 to $-\infty$ to $+\infty$, as usual for linear regression.

Logistic slope coefficients can be interpreted as the effect of a unit of change in the X variable on the predicted logits with the other variables in the model held constant. That is, how a one unit change in X affects the log of the odds when the other variables in the model held constant (Agresti, 2002).

3.3.1.5 Methods of parameter Estimation

Maximum likelihood (ML) estimation is the most common method used to calculate the logit coefficients. This contrasts to the use of ordinary least squares (OLS) estimation of coefficients in linear regression. OLS seeks to minimize the sum of squared distances of the data points to the regression line. ML methods seek to maximize the log likelihood, LL, which reflects how likely the observed values of the outcome may be predicted from the observed values of the predictors.

Suppose $y_1, y_2, \ldots, y_n$ be the *n* independent random observations corresponding to the random variables $Y_1, Y_2, \ldots, Y_n$. Since the Y_i is a Bernoulli random variable, the probability function of Y_i is

$f_i(y_i) = \pi_i^{yi}(1-\pi_i)^{1-yi}$; $y_i = 0$ or 1; i = 1, 2, ..., *n*, since Y"s are assumed to be independent, the joint probability function or likelihood function is given by

$$l(\boldsymbol{\beta}) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{3.7}$$

l is the likelihood function for observed data; it is a function of the parameters. The principle of maximum likelihood states that we use as our estimate of **β** the value which maximizes the expression in *l*. However, it is easier mathematically to work with the log of *l*. The *log likelihood* is defined as

$$L(\boldsymbol{\beta}) = \ln[l(\boldsymbol{\beta})] = \sum_{i=1}^{n} \{y_i \ln(\pi_i) + (1-y_i)\ln(1-\pi_i)\} \tag{3.8}$$

The maximum likelihood estimates are the values of **β** that maximize the above log-likelihood function. Through maximization of the log-likelihood function we can theoretically estimate the vector of parameter **β**. But the equation is nonlinear in **β**, and as a result the estimates do not have a closed form expression. Therefore, **β** will be obtained by maximizing log-likelihood using iterative algorithm method (Agresti, 2002).

3.3.1.6 Goodness of fit of the model

The goodness of fit or calibration of a model measures how well the model describes the response variable. Assessing goodness of fit involves investigating how the predicted values are closer to the observed values.

3.3.1.6.1 The likelihood ratio test (LRT)

The likelihood ratio (LR) test is performed by estimating two models and comparing the fit of one model to the fit of the other. Removing predictor variables from a model will almost always make the model fit worse (i.e., a model will have a lower log likelihood), but it is necessary to test whether the observed difference in model fit is statistically significant. The likelihood ratio test does this by comparing the log likelihoods of the two models. If this difference is statistically significant, then the less restrictive model (the one with more variables) is said to fit the data significantly better than the more restrictive model. If one has the log likelihoods from the models, the likelihood ratio statistic is fairly easy to calculate. The likelihood ratio test is performed to test the overall significance of all coefficients in the model on the basis of the test statistic:

$$G^2 = -2\ln\left(\frac{L(m_1)}{L(m_2)}\right) = 2[LL(m_2) - LL(m_1)] \tag{3.9}$$

where L(m*) denotes the likelihood of the respective model (either model 1 or model 2), and *LL*(m*) is the log likelihood of the respective model. Where m_1 is the more restrictive (reduced) model and m_2 is the less restrictive (saturated) model.

Under the global null hypothesis, H_0: $\beta_1 = \beta_2 = ... = \beta p = 0$, the statistic G^2 follows a chi-square distribution with *p*-1 degrees of freedom and measures how well the independent variables affect the response variable (Hosmer and Lemeshow, 2000).

3.3.1.6.2 The Hosmer–Lemeshow(H-L) test

A better way of assessing the fit of a logistic regression model is comparing the observed and expected numbers of positives for different subgroups of the data. The Hosmer-Lameshow goodness of fit test is useful for assessing overall model fit, particularly when we have many predictor variables, or some of predictor variables are continuous.

The test is similar to a χ^2 goodness of fit test and has the advantage of partitioning the observations into groups of approximately equal size, and therefore there are less likely to be groups with very low observed and expected frequencies. The observations are grouped into g (mostly, g=10) based on the predicted probabilities. For either grouping strategy, the Hosmer-Lemeshow goodness-of-fit statistic, $\hat{C}$, is obtained by calculating the Pearson chi-square statistic from the $g\times 2$ table of observed and estimated expected frequencies. A formula defining the calculation of $\hat{C}$ is as follows:

$$\hat{C} = \sum_{k=1}^{g} \frac{(o_k - n_k'\bar{\pi}_k)^2}{n_k'\bar{\pi}_k(1-\bar{\pi}_k)} \sim \chi^2_{(g-2)} \quad (3.10)$$

where, g denotes the number of groups, n'_k is the total number of observations in the k^{th} group, c_k denotes the number of covariate patterns in the k^{th} decile, o_k is the number of responses among the c_k covariate patterns, and $\bar{\pi}$ is the average estimated probability. The distribution of the statistic $\hat{C}$ is well approximated by the chi-square distribution with $g - 2$ degrees of freedom, $\chi^2_{(g-2)}$ (Hosmer and Lemeshow, 2000).

If p-value for the Hosmer-Lemeshow goodness-of-fit test is greater than 0.05, we will not reject the null hypothesis that there is no difference between observed and model predicted values, implying that the model estimates are adequate to fit the data at an acceptable level or if the observed and expected numbers are sufficiently close, then we can assume that we have an adequate model.

3.3.1.6.3 Testing for the significance of individual predictors

To determine the significance of the predictor variables we can use either the Wald statistic or the likelihood ratio test.

The Wald test

The Wald statistic is an alternative test which is commonly used to test the significance of individual logistic regression coefficients for each predictor variable (that is, to test the null hypothesis in logistic regression that a particular logit coefficient is zero). The Wald test approximates the likelihood ratio test, but with the advantage that it only requires estimating one model.

The Wald statistic is a method which is commonly used to test the significance of individual logistic regression coefficients for each explanatory variable, which is $\beta_i = 0$ vs $\beta_i \neq 0$. For a dichotomous response variable the Wald statistic (W), is the squared ratio of the unstandardized logit coefficient to its standard error, that is

$$W = \left(\frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)} \right)^2 \tag{3.11}$$

The Wald statistic, W, under the null hypothesis is approximately chi-square distributed with 1 degree of freedom. Wald statistics are easy to calculate but their reliability is questionable, particularly for small samples. For data that produce large estimates of the coefficient, the standard error is often inflated, resulting in a lower value of the Wald statistic, and therefore the explanatory variable may be incorrectly assumed to be unimportant in the model. Likelihood ratio tests are generally considered to be superior (Agresti, 2002).

3.3.1.6.4 The ROC Curves

A receiver operating characteristics (ROC) curve is a plot of sensitivity (True positive) as a function of (1-specificity) for the all possible cutoffs π_0. The ROC curve is more informative than the classification table, since it summarizes predictive power for all possible π_0. First, we calculate sensitivity and specificity pairs for each possible cutoff point and plot sensitivity on the y axis by (1-specificity) on the x axis. When π_0 gets near 0, almost all predictions are $\hat{y} = 1$; then, sensitivity is near 1, specificity is near 0, and the point for (1 – specificity, sensitivity) has coordinates near (1, 1). When π_0 gets near 1, almost all predictions are $\hat{y} = 0$; then, sensitivity is near 0, specificity is near 1, and the point for (1 –

specificity, sensitivity) has coordinates near (0, 0). The ROC curve usually has a concave shape connecting the points (0, 0) and (1, 1). The area under the ROC curve, which ranges from zero to one, provides a measure of the model's ability to *discriminate* between those subjects who experience the outcome of interest versus those who do not.

Hosmer and Lemeshow provide general rules for interpreting area under ROC curve values. Paraphrasing their rules give the general guidelines below:

ROC = 0.5	No discrimination
0.7≤**ROC** < 0.8	Acceptable discrimination
0.8 ≤**ROC** <0.9	Excellent discrimination
ROC ≥ 0.9	Outstanding discrimination

3.3.1.7 Model diagnostics

3.3.1.7.1 Outlier detection

An outlier is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism. An observation is influential if it is individually or together with several other observations, has a demonstrably larger impact on the calculated values of various estimates than is the case for most of the other observations.

Detecting outliers is common practice and it is important because outliers can affect the regression model in two ways: outliers may almost uniquely determine regression coefficients; they may also cause the standard errors of regression coefficients to be much smaller than they would be if the observation were excluded. There are two types of outliers, so that it is important to distinguish between the two types. Outliers in the response variable represent model failure. Outliers with respect to the predictors are called leverage points (Vittinghoff *et al*, 2005).

Leverage Value is a term used in connection with regression analysis and, in particular, in analyses aimed at identifying those observations which have a large effect on the outcome of fitting regression models. Leverage points are those observations, if any, made at outlying values of the predictor variables such that the lack of neighboring observations means that the fitted regression model will pass close to that particular observation. Leverages are obtained from the diagonal element of the hat matrix, H, which is given as

$$H = V^{1/2} X (X'VX)^{-1} X' V^{1/2} \tag{3.12}$$

where V is the $j \times j$ diagonal matrix with elements $m_j \hat{\pi}_j (1 - \hat{\pi}_j)$ and X is $j \times (p+1)$ design matrix.

Let the quantity h_j denote the j^{th} diagonal element of the $j \times j$ hat matrix H defined in equation (3.12). It may be shown that

$$h_j = m_j \hat{\pi}_j (1 - \hat{\pi}_j) \mathbf{x'}_j (\mathbf{X'VX})^{-1} \mathbf{x'}_j = v_j \times b_j \tag{3.13}$$

where $b_j = \mathbf{x'}_j (\mathbf{X'VX})^{-1} \mathbf{x'}_j$ and $\mathbf{x'}_j = (1, x_{1j}, x_{2j}, \ldots, x_{pj})$ is the vector of covariate values defining the j^{th} covariate pattern.

The greater the value of h_j (i.e. $h_j > 1$), the more potential that observation has for influencing the model fit.

3.3.1.7.2 Influential Statistics

Influential values are points that have exerted excessive influence on the regression coefficient estimates. Influential measures can be used to identify cases that are highly influential on the logistic regression estimates. An influential point affects the statistical significance as well as the strength and direction of the association between a response variable and predictor variables. The two common measures of the influence of an observation are Cook‟s distance and DFBETAS.

Cook's Distance (D) is a measure of the influence of excluding any specific observation on the set of parameter estimates. D is greater than 0, and may be arbitrarily large. The value of D>1 identifies cases that might be influential (Cook, 1977). Cook‟s D_i, i= 1, 2,…, n statistic is defined as:

$$D_i = \frac{(\hat{\beta}_{(-i)} - \hat{\beta})' X' X (\hat{\beta}_{(-i)} - \hat{\beta})}{ps^2} \tag{3.14}$$

where $\hat{\beta}_{(-i)}$ denote the least squares estimate of β with the i^{th} observation deleted.

DFBETAS: For each parameter estimate, a DFBETAS diagnostic is calculated for each observation. This is the standardized difference in the parameter estimate due to deleting the observation, and it can be used to assess the effect of an individual observation on each estimated parameter of the fitted model. These measures are useful for detecting observations that are causing instability in the selected coefficients.

The influential observations for the individual regression coefficients are identified by $DFBETAS_{j(i)}$, j = 0, 1, 2, . . . , p and calculated as

$$DFBETAS_{j(i)} = \frac{\hat{\beta}_j - \hat{\beta}_{j(i)}}{S_i \sqrt{c_{jj}}} \tag{3.15}$$

Where c_{jj} is the j^{th} diagonal element of $(X'X)^{-1}$. Values of $DFBETAS_{j(i)}$ larger than 1 indicate the considerable influence an observation i has on the j^{th} regression coefficient.

3.3.1.8 Model Building Strategy (Variable Selection)

In modeling with many explanatory variables, one is usually concerned with the goal of selecting those variables that result in the "best" model within the scientific context of the problem. Having a basic plan to follow in selecting the variables for the model and assessing the adequacy of the model both in terms of the individual variables and from the point of view of the overall fit of the model is required for achieving this "best" model. Successful modeling of a complex data set is part science, part statistical methods, and part experience and common sense (Hosmer and Lemeshow, 2000).

The criteria for including a variable in a model may vary from one problem to the next and from one scientific discipline to another. The traditional approach to statistical model building involves seeking the most parsimonious model that still explains the data. The rationale for minimizing the number of variables in the model is that the resultant model is more likely to be numerically stable, and is more easily generalized. The more variables included in a model, the greater the estimated standard errors become, and the more dependent the model becomes on the observed data. Epidemiologic methodologists suggest including all clinically and intuitively relevant variables in the model, regardless of their "statistical significance."

There are several steps one can follow to aid in the selection of variables for a logistic regression model. The process of model building is quite similar to the one used in multiple linear regression.

(1) The selection process should begin with a careful univariable analysis of each variable. For nominal, ordinal, and continuous variables with few integer values, we suggest this be done with a contingency table of outcome ($y = 0, 1$) versus the p levels of the independent variable.

(2) Upon completion of the univariable analysis, we select variables for the multivariable analysis. Any variable whose univariable test has a p-value < 0.25 is a candidate for the multivariable model along with all variables of known clinical importance. Once the variables have been identified, we begin with a model containing all of the selected variables.

Another approach to variable selection is to use a stepwise method in which variables are selected either for inclusion or exclusion from the model in a sequential fashion based solely on statistical criteria. There are two main versions of the stepwise procedure:

(a) Forward selection with a test for backward elimination and

(b) Backward elimination followed by a test for forward selection. The algorithms used to define these procedures in logistic regression.

Forward selection adds terms sequentially until further additions do not improve the fit. At each stage it selects the term giving the greatest improvement in fit.

Backward elimination begins with a complex model and sequentially removes terms. At each stage, it selects the term for which its removal has the least damaging effect on the model (e.g., largest *P*-value). The process stops when any further deletion leads to a significantly poorer fit.

We can also use information criteria for model selection. The basic idea behind the information criteria is penalizing the likelihood for the model complexity-the number of explanatory variables used in the model. The most popular in this family are the Akaike information criterion (AIC) and Schwarz information criterion (SC). The AIC and SC can be defined by the equations:

$$AIC = -2\log L(M) + 2p \tag{3.16}$$

$$SC = -2\log L(M) + p\log(n) \tag{3.17}$$

where log L(M) is the maximized log likelihood for fitted model,

p = number of parameters in the model,

n = Sample size

N.B: The lower the values of AIC, SC the best fit of the model.

CHAPTER FOUR

RESULTS AND DISCUSSION

This chapter presents the results and discussion of this study. The aim of the study was to identify the determinants of neonatal mortality in Ethiopia. Descriptive, bivariate and multivariable statistical analyses will be used to fulfill the aim of the study. Simple logistic regression analysis was employed to see the association of each predictor with neonatal mortality separately. Thereafter, the multiple logistic regression analysis has been used to see the net effect of the predictor variables on neonatal mortality.

4.1 Results of descriptive analysis

A total of 6,303 neonates born within five years preceding the 2011 EDHS whose complete information was available in the survey were included in this study. Out of these, 378 (6%) neonates died. Table 4.1 presents percentage distribution of neonates in Ethiopia by background characteristics. Of the total of 6,303 neonates, 51.0% (49.0%) were males (females); 15.8% (84.2%) were born in urban (rural) areas; 97.5% (2.5%) were single (multiple); 81.1% (11.9%) were born at home (health facility). High percentages of neonates were born to mothers residing in Oromiya region (15.0%) followed by mothers residing in SNNP (14.5%) while the lowest figure was in Addis Ababa (2.8%). With regard to educational level, 73.0% mothers & 54.5% fathers had no education while 23.3% mothers and 36.0% fathers had primary education and the remaining 3.7% mothers and 9.5% fathers had secondary and higher education. About 49.8%, 16.6% and 33.6% households were classified as poor, medium income and rich, respectively. More than half (53.1%) of neonates were born to mothers of Christianity followers while 44.9% neonates were born to mothers of Muslim followers and the rest 2.0% neonates were born to mothers who follow Traditional/Other beliefs.

There were also 1.9%, 68.0% and 30.0% live-births born to mothers in the age group of 15-19, 20-34 and 35-49, respectively. 54.7% neonates were 1-4 ranked birth while 45.3% were 5 or more ranked birth. 18.1% neonates were born within the interval of <24 months while 35.8% neonates were born within the interval of 24-36 months and the rest 46.1% neonates were born within the interval of > 36 months of the preceding birth. 32.2% neonates had small size while 38.9% had average size and the rest 29.0% had large size at birth. 58.0% mothers didn‟t visit ANC during pregnancy while 22.8% and 19.2% mothers visited ANC of < 4 and ≥ 4 during pregnancy, respectively.

Table 4.1: Percentage distribution of neonates in Ethiopia by background characteristics, EDHS 2011

Variable	Category	Neonatal death		Total
		No	Yes	
Residence	Urban	944	54	998 (15.8%)
	Rural	4981	324	5,305 (84.2%)
Region	Tigray	647	41	688 (10.9%)
	Affar	573	23	596 (9.5%)
	Amhara	740	53	793 (12.6%)
	Oromiya	887	59	946 (15.0%)
	Somali	468	31	499 (7.9%)
	Ben-Gumuz	479	45	524 (8.3%)
	SNNP	863	50	913 (14.5%)
	Gambela	444	34	478 (7.6%)
	Harari	308	25	333 (5.3%)
	Dire Dawa	347	10	357 (5.7%)
	Addis Ababa	169	7	176 (2.8%)
Mother‟s education level	No Education	4337	262	4,599 (73.0%)
	Primary	1369	100	1,469 (23.3%)
	Secondary &Higher	219	16	235 (3.7%)
Father‟s education level	No Education	3243	190	3,433 (54.5%)
	Primary	2128	144	2,272 (36.0%)
	Secondary & Higher	554	44	598 (9.5%)

House hold"s Wealth index	Poor	2934	204	3,138 (49.8%)
	Medium	986	59	1,045 (16.6%)
	Rich	2005	115	2,120 (33.6%)
Religion	Christian	3142	206	3,348 (53.1%)
	Muslim	2665	162	2,827 (44.9%)
	Traditional/Other	118	10	128 (2.0%)
Mother"s age at Child birth	15-19	86	36	122 (1.9%)
	20-34	4036	252	4,288 (68.0%)
	35-49	1803	90	1,893 (30.0%)
Birth order number	1-4	3198	249	3,447 (54.7%)
	5 or more	2727	129	2,856 (45.3%)
Preceding birth Interval (month)	<24	1028	113	1,141 (18.1%)
	24-36	2123	134	2,257 (35.8%)
	>36	2774	131	2,905 (46.1%)
Sex	Male	2998	214	3212 (51.0%)
	Female	2927	164	3,091(49.0%)
Child is Twin	Single	5827	319	6,146 (97.5%)
	Multiple	98	59	157 (2.5%)
Size of child at birth	Average	2337	113	2,450 (38.9%)
	Small	1916	111	2,027 (32.2%)
	Large	1672	154	1,826 (29.0%)
Number of Antenatal visits	None	3464	191	3,655 (58.0%)
	<4	1337	98	1,435 (22.8%)
	≥ 4	1124	89	1,213 (19.2%)
Place of delivery	Home	5241	314	5,555 (88.1%)
	Health facility	684	64	748 (11.9%)

4.2 Results of bivariate logistic regression analysis

Bivariate (simple) logistic regression was used to examine the individual effects of each of the selected predictors on neonatal mortality. The results given in Table 4.2 show that predictors Region, Maternal age, Birth order, Birth interval, Sex, Type of birth, Birth size, ANC and Place of delivery were significantly associated with neonatal mortality at 5% level of significance. That is, the individual contribution of each of these predictors to neonatal mortality is significant.

The covariates Residence, Religion, Wealth index, Maternal education and Paternal education are not significantly associated with neonatal mortality at 5% level of significance in this bivariate (simple) logistic regression analysis. By assessing their univaraite tests, the predictors Wealth index (p-value =0.238) and paternal education (p-value=0.156) were selected as candidate to be included in multiple logistic regression analysis at 25% level of significance.

Hence, on the basis of the univariate results, the predictors that were considered as candidates for multiple logistic regression analysis were Region, Paternal education, Wealth index, Maternal age, Birth order, Birth interval, Sex, Type of birth, Birth size, ANC and Place of delivery.

Table 4.2: Results of bivariate logistic regression analysis of neonatal mortality for selected covariates, EDHS 2011

Variables	df	Estimate	S.E	Wald	Sig.	OR	95% CI	
							lower	Upper
Residence	**1**			**0.722**	**0.395**			
Urban(R)								
Rural)	1	-0.128	0.151	0.722	0.395	0.879	0.654	1.183
Region	**10**			**21.612**	**0.017***			
Tigray	1	0.425	0.418	1.035	0.309	1.530	0.674	3.471
Affar	1	-0.031	0.440	0.005	0.943	0.969	0.409	2.298
Amhara	1	0.548	0.411	1.775	0.183	1.729	0.773	3.870
Oromiya	1	0.474	0.408	1.345	0.246	1.606	0.721	3.576
Somali	1	0.470	0.428	1.203	0.273	1.599	0.691	3.700
Ben-Gumuz	1	0.819	0.416	3.875	0.049*	2.268	1.004	5.126
SNNP	1	0.336	0.412	0.663	0.416	1.399	0.624	3.138
Gambela	1	0.615	0.425	2.093	0.148	1.849	0.804	4.251
Harari	1	0.673	0.438	2.357	0.125	1.960	0.830	4.626
Dire Dawa	1	-0.363	0.502	0.523	0.470	0.696	0.260	1.860
Addis Ababa(R)								
MOTHEDUC	**2**			**2.713**	**0.258**			
No education(R)								
Primary	1	0.063	0.113	0.316	0.574	1.209	0.953	1.535
2ndary & Higher	1	0.064	0.177	0.128	0.721	1.209	0.717	2.040
PEDUC	**2**			**3.720**	**0.156**			
No education(R)								
Primary	1	-0.304	0.174	3.075	0.080	0.738	0.525	1.036
2ndary & Higher	1	-0.160	0.179	0.803	0.370	0.852	0.600	1.209
Wealth, Rich(R)	**2**			**2.868**	**0.238**			
Poor	1	0.192	0.120	2.566	0.109	1.212	0.958	1.534

Medium	1	0.042	0.165	0.066	0.797	1.043	0.755	1.441
Religion	**2**			**1.243**	**0.537**			
Christian(R)								
Muslim	1	-0.257	0.337	0.580	0.446	0.774	0.400	1.498
Traditional/other	1	-0.332	0.339	0.960	0.327	0.717	0.369	1.394
AGE, 35-49(R)	**2**			**93.739**	**0.000***			
15-19	1	2.127	0.226	88.548	0.000*	8.386	5.385	13.059
20-34	1	0.224	0.126	3.154	0.076	1.251	0.977	1.601
Bord	**1**			**19.950**	**0.000***			
1-4	1	0.498	0.112	19.950	0.000*	1.646	1.323	2.048
5 or more(R)								
Pre_interval	**2**			**40.706**	**0.000***			
< 24	1	0.845	0.133	40.064	0.000*	2.328	1.792	3.024
24-36	1	0.290	0.126	5.284	0.022*	1.337	1.044	1.712
>36 (R)								
Sex, Female(R)	**1**			**5.123**	**0.024***			
Male	1	0.242	0.107	5.123	0.024*	1.274	1.033	1.571
Birth_Type	**1**			**188.730**	**0.000***			
Single	1	-2.398	0.175	188.730	0.000*	0.091	0.065	0.128
Multiple(R)								
Birth_size	**2**			**27.880**	**0.000***			
Average (R)								
Small	1	-0.644	0.128	25.369	0.000*	0.525	0.409	0.675
Large	1	-0.464	0.129	12.932	0.000*	0.629	0.489	0.810
ANC, ≥ 4(R)	**2**			**9.413**	**0.009***			
None	1	-0.362	0.133	7.420	0.006*	1.696	0.537	0.903
<4	1	-0.077	0.152	0.258	0.611	0.926	0.687	1.247
POD, Home (R)	**1**			**9.712**	**0.002***			
Health facility	1	-0.446	0.143	9.712	0.002*	0.640	0.484	0.848

* Significant at 5% level of significance   OR: Crude Odd Ratio R: reference category

4.3 Results of multiple logistic regression analysis

Based on the results of bivariate logistic regression analysis, 11 selected predictors were included in the multiple logistic regression analysis to assess their net effect. Stepwise method of variable selection was employed using SAS software.

Out of the 11 predictors considered in this section, 9 predictors were significantly associated with neonatal mortality in Ethiopia. The predictors Region, wealth index, Maternal age, Birth order, birth interval, Type of birth, Birth size, ANC and Place of delivery were found to have significant net effect on neonatal mortality. The remaining 2 predictors were not significantly associated with neonatal mortality. These predictors are sex of child and paternal education level. Sex of child which was significant in the bivariate logistic regression analysis lost its significant association with neonatal mortality in the multiple logistic regression analysis. Paternal education is still not significantly associated with neonatal mortality at 5% level of significance as in the case of bivariate logistic regression analysis. On the other hand, household‟s wealth index which was insignificant in the bivariate logistic regression analysis is found to be significantly associated with neonatal mortality in the multiple logistic regression analysis. Table 4.3 presents results of multiple logistic regression analysis for factors found to be associated with neonatal mortality. The adjusted odds ratios (AOR), 95% confidence interval (CI), p-value and Wald chi-square test statistic of each factor fitted into the model were shown in this table.

Compared to neonates born in Addis Ababa, neonates born in Amhara region are 3.120 times more likely to die (AOR=3.120; 95% CI: 1.288-7.558) and neonates born in Benishangul-Gumuz region are 3.039 times more likely to die (AOR=3.039; 95% CI: 1.242-7.438). However, neonates born in Dire Dawa are 0.888 times less likely to die (AOR=0.888; 95% CI: 0.319-2.469) compared to neonates born in Addis Ababa. Controlling for other variables in the model neonates of poor households had 51.7% increased risk (AOR =1.517; 95% CI: 1.139-2.021) of neonatal mortality compared to neonates of rich households.

In terms of maternal age at child birth, neonates born to mothers in the age group 15-19 are 4.106 times more likely to die (AOR=4.106; 95% CI: 2.412-6.987) compared to neonates born to mothers in the age group 35-49 while neonates born to mothers in age group 20-34 are 0.841 times less likely to die (AOR=0.841; 95% CI: 0.620-1.141). Compared to 5 or more ranked neonates, the 1-4 ranked neonates had 1.751 fold higher risk (AOR=1.751; 95% CI: 1.318-2.326) of neonatal mortality. Single births are 0.075 times less likely to die (AOR =0.075; 95% CI: 0.051-0.109) compared to multiple births.

The results also show that neonates born within the interval of < 24 months are nearly 2.7 times more likely to die (AOR=2.666; 95% CI: 2.003-2.549) compared to neonates born within the preceding birth interval of >36 months. Controlling for other variables in the model, large size neonates are 85.4% more likely to die (AOR=1.854; 95% CI: 1.423-2.415) compared to average size neonate at birth. Neonates born to mothers who had no ANC visits during pregnancy had 62.4% higher risk (AOR=1.624; 95% CI: 0.458-0.851) of neonatal mortality compared to neonates born to mothers who had ANC of ≥ 4 during pregnancy.

Contrary to the widely held expectation, deliveries at health facility had 1.790 times higher risk (AOR=1.790; 95%CI: 1.249-2.564) of neonatal mortality in Ethiopia compared to deliveries at home. However, this result should be interpreted with caution because most deliveries start at home and are referred to health facilities as a result of complications in Ethiopia e.g. prolonged labour, intra-partum haemorrhage. Thus, deliveries at heath facilities are more complicated and more likely to have poor outcomes.

Table 4.3: Results of multiple logistic regression analysis for factors found to be associated with neonatal mortality, EDHS 2011

Variables	DF	Est.($\hat{\beta}$)	S.E	Wald χ^2	P-value	AOR	95% Wald CLs	
Intercept	1	-0.9968	0.1369	52.9941	<.0001*			
Region	**10**			**25.0962**	**0.0052***			
Tigray	1	0.1072	0.1712	0.3919	0.5313	2.089	0.855	5.104
Affar	1	-0.3379	0.2192	2.3753	0.1233	1.339	0.517	3.466
Amhara	1	0.5083	0.1543	10.8497	0.0010*	3.120	1.288	7.558
Oromiya	1	0.2020	0.1466	1.8974	0.1684	2.297	0.956	5.521
Somali	1	0.0797	0.1937	0.1692	0.6808	2.032	0.808	5.115
Ben-Gumuz	1	0.4820	0.1664	8.3931	0.0038*	3.039	1.242	7.438
SNNP	1	0.1591	0.1572	1.0234	0.3117	2.200	0.903	5.361
Gambela	1	0.0396	0.1919	0.0426	0.8365	1.953	0.785	4.858
Harari	1	0.1383	0.2182	0.4017	0.5262	2.155	0.864	5.378
Dire Dawa	1	-0.7486	0.3048	6.0338	0.0140*	0.888	0.319	2.469
A. A(R)								
Wealth	**2**			**8.2065**	**0.0165***			
Poor	1	0.1974	0.0829	5.6715	0.0172*	1.517	1.139	2.021
Medium	1	0.0223	0.1055	0.0445	0.8329	1.274	0.887	1.829
Rich(R)								
AGE	**2**			**48.3758**	**<.0001***			
15-19	1	0.9991	0.1592	39.3652	<.0001*	4.106	2.412	6.987
20-34	1	-0.5859	0.0943	38.5827	<.0001*	0.841	0.620	1.141
35-49 (R)								
Bord	**1**			**14.9662**	**0.0001***			
1-4	1	0.2801	0.0724	14.9662	0.0001*	1.751	1.318	2.326
5 or more (R)								
Pre_interval	**2**			**45.2869**	**<.0001***			

<24	1	0.5165	0.0843	37.5178	<.0001*	2.666	2.003	3.549
24-36	1	-0.0524	0.0784	0.4468	0.5039	1.509	1.157	1.969
>36 (R)								
Birth_Type	**1**			**181.7021**	**<.0001***			
Single	1	-1.2972	0.0962	181.7021	<.0001*	0.075	0.051	0.109
Multiple (R)								
Birth_size	**2**			**24.6491**	**<.0001***			
Average (R)								
Small	1	-0.1475	0.0844	3.0526	0.0806	1.091	0.820	1.452
Large	1	0.3824	0.0782	23.9378	<.0001*	1.854	1.423	2.415
ANC	**2**			**9.6909**	**0.0079***			
None	1	-0.2457	0.0817	9.0497	0.0026*	1.624	0.458	0.851
<4	1	0.0206	0.0881	0.0547	0.8151	0.815	0.586	1.133
≥ 4 (R)								
POD	**1**			**10.0672**	**0.0015***			
Home (R)								
Health facility	1	0.2911	0.0917	10.0672	0.0015*	1.790	1.249	2.564

***significant at 5% at level of significance   AOR: Adjusted Odd Ratio R: reference category

For the sake of comparison of the results of bivariate analysis and multiple logistic regression analysis; predictors Region, Maternal age, Birth order, birth interval, Type of birth, Birth size, ANC and Place of delivery which were significantly associated with neonatal mortality at 5% level of significance in the bivariate analysis still hold their significant association with neonatal mortality in the multiple logistic regression analysis. Sex of neonate which was significantly associated with neonatal mortality in the bivariate analysis lost its significant association with neonatal mortality in the multiple logistic regressions analysis. On the other hand, household‟s wealth index which was not significantly associated with neonatal mortality in the bivariate analysis is found to be significantly associated with neonatal mortality in the multiple logistic regressions analysis.

4.3.1 Goodness of fit of the model

In this section goodness of fit of the model has been assessed. The model fit statistics are given in Table 4.4 below. The values of AIC, SC and -2Log L given in this table indicate that the model with intercept and covariates has smaller values for these criteria compared to the model with intercept only. This indicates that the model fitted with the covariates is better than the model with no covariates. That is, the selected covariates are important determinants of neonatal mortality in Ethiopia.

Table 4.4: Model Fit Statistics

Criterion	Intercept only	Intercept & Covariates
AIC	2862.162	2569.917
SC	2868.911	2731.888
-2 Log L	2860.162	2521.917

We can also test the goodness of the fitted model by using likelihood ratio, score and Wald tests. The chi-square, degree of freedom and the corresponding p-values of these tests are given in Table 4.5. All of these tests are significant at 5% level of significance indicating the adequacy of the model.

Table 4.5: Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	338.2447	23	<.0001
Score	520.8018	23	<.0001
Wald	339.8581	23	<.0001

The overall goodness of fit of the model can also be assessed using the Hosmer-Lemeshow goodness-of-fit test. The value of the Hosmer-Lemeshow goodness-of-fit statistic and the corresponding p-value computed from the chi-square distribution with 8 degrees of freedom are given in Table 4.6 below. Since the p-value =0.1568 > 0.05 at 5% level of significance, we don‟t reject the null hypothesis of no difference between observed and model predicted values and conclude that the model is adequate.

Table 4.6: Hosmer-Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
11.8764	8	0.1568

Another way of testing the goodness of fit of the model is by using ROC curve**.** The area under the ROC curve, which ranges from zero to one, provides a measure of the model ability to *discriminate* between those neonates who died versus those who survived. The area under the ROC curve is 0.744 which lies between 0.7 and 0.8. This indicates acceptable discrimination according to Hosmer-Lemeshow rule. The discrimination is statistically significant with p-value=0.000 (Table 4.7). Therefore, the fitted model acceptably discriminates between those neonates who died and who do not.

Table 4.7: Area Under ROC Curve

Area	S.E[a]	Asymptotic Sig.[b]	Asymptotic 95% CI	
			Lower Bound	Upper Bound
0.744	0.014	0.000	0.718	0.771

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

4.3.2 Model Diagnostics

The adequacy of the fitted model was checked for possible presence and treatment of outliers and influential values. The diagnostic test results for detection of outliers and influential values are presented in Table 1B under Appendix 1. The DFBETAs for model parameters including the constant term, Cook"s influence statistic and leverage values were less than unity. DFBETAs less than unity imply that no specific impact an observation had on the coefficient of a particular predictor variable. Cook"s distance less than unity shows that an observation had no overall impact on the estimated vector of regression coefficients ($\hat{\boldsymbol{\beta}}$). The value of the leverage statistic less than unity shows that no subject has a substantially large impact on the predicted values of model.

4.4 Discussion

The main objective of this study was to identify socio-economic, demographic, nutritional and maternal health care seeking determinants of neonatal mortality in Ethiopia using the nationally representative 2011 EDHS data. A total of 6,303 neonates born within five years preceding the 2011 EDHS whose complete information was available in the survey were included in the study. Out of these neonates 378(6%) died. Descriptive, bivariate and multiple logistic regression statistical methods of data analyses were employed to identify determinants of neonatal mortality in Ethiopia.

Bivariate analysis was used to assess the association between neonatal mortality and independent variables. The results revealed that predictors Region, maternal age, Birth order, birth interval, Sex, Type of birth, Birth size, ANC and Place of delivery were significantly associated with neonatal mortality at 5% level of significance. The covariates Residence, Religion, Wealth index, Maternal education and Paternal education were not significantly associated with neonatal mortality at 5% level of significance in bivariate analysis. The results of multiple logistic regression analysis showed that Region, Maternal age, Birth order, birth interval, Type of birth, Birth size, ANC and Place of delivery still hold their significant association with neonatal mortality as in the case of bivariate analysis. Additionally, household"s wealth index was found to be significantly associated with neonatal mortality in the multiple logistic regression analysis.

The findings of the study revealed that most of proximate determinants had significant association with neonatal mortality than socio-economic factors in Ethiopia. Proximate determinants maternal age, Birth order, birth interval, Type of birth, Birth size, ANC and Place of delivery were significantly associated with neonatal mortality in Ethiopia. Socio-economic factors, region and household"s wealth index had significant association with neonatal mortality in Ethiopia.

It was found that neonates born in Amhara and Benishangul-Gumuz regions were more likely to die compared to neonates born in Addis Ababa. This could be due to differences in socio-economic and functioning of the healthcare system. Further investigation is needed to identify specific factors that are relevant to these regions. Identifying such factors will help programmers to design and implement region-specific interventions to better address neonatal mortality. The result is consistent with a study conducted by Omariba *et al*., (2008) in Kenya. Consistent with the study conducted by Ikamari (2013) in Kenya, the results of this study showed that neonates belonging to poor households had a higher risk of neonatal mortality compared to neonates belonging to rich household.

Studies by Markovitz et al. (2005) and Seedhom and Kamal (2010) found higher risk of neonatal mortality among neonates born to younger adolescent mothers than older. The findings of this study showed significant increased risk of neonatal mortality among neonates born to mothers in the age group (15-19) compared to neonates born to mothers in the age group 35-49. This is partially explained by differences in socioeconomic factors in younger versus older women and is mediated primarily through preterm delivery, small for gestational age (SGA) and low birth weight (LBW) or some interaction of these variables.

Birth order of a child showed controversial results for neonatal mortality. Some studies showed that first or lower ranked births were at higher risk of neonatal mortality while others showed that higher ranked births were at increased risk of neonatal mortality. For instance, a study conducted in Taiwan (Wang and Lin, 1999) showed that children with first and fifth ranked births were at higher risk of early neonatal deaths. Hailemariam and Tesfaye (1997) showed that children with sixth or higher order births were at increased risk of neonatal mortality in Ethiopia. It is argued that, in most of the developing countries, higher mortality risks were found for the first-born children compared to birth order two through six (Hobcraft et al., 1985). Our study showed that (1-4) ranked births were significantly at increased risk of neonatal mortality compared to 5 or more ranked births.

The results of this study showed higher odds of neonatal mortality for babies born within the interval of <24 months compared to babies born within the interval of >36 months of the preceding birth. The finding is consistent with a previous study by Boerma and Bicego (1992) where it was shown that short preceding birth intervals are associated with an increased risk of dying in the neonatal period and at 1-6 months of age, and to a much lesser extent at 7-23 months of age.

Multiple births are relatively rare events, but contribute substantially to mortality in both neonatal and post-neonatal periods (Alam et al., 2007). The result of the present study showed a reduced risk of neonatal mortality for single births compared to multiples. This may be due to the fact that multiple births are more likely associated with low birth weight and biological immaturity.

Size of child at birth has a significant net effect on neonatal mortality in Ethiopia. The findings of our study showed that large size child at birth had a significantly higher risk of neonatal mortality compared to average size neonate at birth. This finding is supported by the study conducted in Pakistan by Mahmood (2002) that has found babies of large size had more than 8.5 times higher risk of neonatal mortality compared to average (normal-size) born babies.

ANC seeking is another important determinant of neonatal mortality. The result of this study revealed that the risk of neonatal mortality was significantly higher among neonates whose mothers did not seek ANC services during pregnancy. The antenatal period clearly presents opportunities for reaching pregnant women with a number of interventions that may be vital to their health and well-being and that of their infants (Abou-Zahr and Wardlaw, 2003). The putative benefits of ANC for babies include increased growth, reduced risk of infection and increased survival (Campbell and Graham, 2006). Our findings are consistent with studies conducted by (Abou-Zahr and Wardlaw, 2003; Campbell and Graham, 2006).

Consistent with the study conducted in Swaziland by Zwane and Masango (2012), the results of this study indicated that neonates delivered at health facilities had significantly higher risk of neonatal mortality compared to neonates delivered at home. This may be due to most deliveries start at home and are referred to health facilities as a result of complications, e.g. prolonged labour, intra-partum haemorrhage. Thus, deliveries at health facilities become more complicated and more likely to have poor outcomes.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The study aimed to identify the determinants of neonatal mortality in Ethiopia using the nationally representative 2011 EDHS data. The results of descriptive analysis showed that of the 6,303 neonates included in this study, 378(6%) died within 28 days after birth. The results of multiple logistic regression analysis showed that the predictors Region, Wealth index, maternal age, Birth order, Birth interval, Type of birth, Birth size, Antenatal care and Place of delivery had significant influence on neonatal mortality in Ethiopia. Hence, public health interventions directed at reducing neonatal mortality should give special attention to these predictors for further reduction of neonatal mortality in Ethiopia. Furthermore, short preceding birth interval and child-bearing at early age have highly significant influence on neonatal mortality in Ethiopia.

The results also indicated that most of the proximate determinants Maternal age, Birth order, birth interval, Type of birth, Birth size, Antenatal care and Place of delivery are significantly associated with neonatal mortality in Ethiopia. Region and Wealth are the only socio-economic factors significantly associated with neonatal mortality in Ethiopia.

5.2 Recommendations

Based on the results of the study the following recommendations are forwarded.

- ❖ The study recommends that concerted efforts should be made to reduce the neonatal mortality in the high mortality regions, particularly in Amhara and Benishangul-Gumuz regions.
- ❖ Since household"s wealth index is an important determinant of neonatal mortality in Ethiopia, concerted efforts should be made to improve the economic status of the households particularly that of poor household"s wealth index.
- ❖ Efforts should also be made to increase birth intervals since proper spacing of births would provide more time for childcare by making more maternal resources available for the care of the child and mother.
- ❖ It is also recommended that efforts should be made to reduce child-bearing at early ages.
- ❖ We also recommend that concerted efforts should be made to encourage the uptake of antenatal care during pregnancy.
- ❖ Short birth interval neonates, child-bearing at early ages and antenatal care utilization should be taken into account when planning the interventions to reduce neonatal mortality in Ethiopia.
- ❖ We also recommend that efforts should be made to have vital registration systems that are complete enough to provide accurate estimates of neonatal mortality in Ethiopia.
- ❖ Finally we recommend that further research should be conducted to identify determinants of neonatal mortality in Ethiopia using other predictors.

REFERENCES

1. Abou-Zahr, C.L. & T. Wardlaw (2003). Antenatal care in developing countries: Promises, achievements and missed opportunities—An analysis of trends, levels and differentials 1990–2001. Geneva: World Health Organization.
2. Agresti, A. (2002). An Introduction to Categorical data Analysis. 2nd Edition, John Wiley and Sons. Inc., New York.
3. Alam N, Van Ginneken JK, Bosch AM. Infant mortalityamong twins and triplets in rural Bangladesh in 1975-2002. *Trop Med Int Health* 2007; 12:1506-14.
4. Andersson et al. The impact of neonatal mortality on subsequent survival in rural Ethiopia. *Annals of Tropical Paediatrics (2002) 22, 25-32.*
5. Araújo BF, Bozzetti MC, Ana C.A. Tanaka A.C.A (2000). Early neonatal mortality in Caxias do Sul:a cohort study. *J Pediatr (Rio J) 2000; 76(3): 200-6.*
6. Black RE, Morris SS, Bryce J. 2003. Where and why are 10 million children dying every year? *The Lancet* 361(9376):2226-2234.
7. Boerma, J. T. and G. T. Bicego. (1992). "Preceding Birth Intervals and Child Survival: Searching for Pathways of Influence." Studies in Family Planning, 23(4): 243 – 56.
8. Campbell, O.M.R. & W.J. Graham (2006). Strategies for reducing maternal mortality: Getting on with what works. *Lancet*, 368(9543), 1284–99.
9. Central Statistical Agency [Ethiopia] and ICF International. 2012. Ethiopia Demographic and Health Survey 2011. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International.
10. Central statistical Agency [Ethiopia] and ORC Macro (2006). Ethiopia Demographic and Health Survey 2005. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ORC Macro.
11. *Chaman* et al. Neonatal Mortality Risk Factors in a Rural Part of Iran: A Nested Case-Control Study. *Iranian J Publ Health 2009, Vol. 38, No.1, pp. 48-52.*
12. Chowdhury AH, Islam.S.M.S, Karim A (2013). Covariates of Neonatal and Post-Neonatal Mortality in Bangladesh. Global Journals Inc. (USA).
13. Chowdhury QH, Islam R and Hossain K. Socio-economic determinants of neonatal, post neonatal, infant and child mortality. International Journal of Sociology and AnthropologyVol.2 (6), pp.118-125, June 2010.
14. Cook, R. D. (1977). Detection of influential observations in linear regression. Technometrics, 19, 15-18.
15. Diallo A. H., Meda N., Ouédraogo W. T., Cousens S. and Tylleskar T. (2011). A prospective study on neonatal mortality and its predictors in a rural area in Burkina Faso: Can MDG-4 be met by 2015. Journal of Perinatology 31, 656–663.

16. Habshah Midi, S.K. Sarkar & Sohel Rana (2010) Collinearity diagnostics of binary logistic regression model, *Journal of Interdisciplinary Mathematics, 13:3, 253-267*, DOI: 10.1080/09720502.2010.10700699.
17. Hailemariam A, Tesfaye M. Determinants of infant and early childhood mortality in a small urban community of Ethiopia: a hazard model analysis. *Ethiopian J Health Dev* 1997; 11:189-200.
18. Haines A, Cassels A. Can the millennium development goals be attained? *BMJ* 2004; 329:394-97.
19. Healy LM (2006). An overview of Logistic Regression: Eastern Michigan University, College of Technology.
20. Hobcraft JN, Mcdonald JW, Rutstein SO. Demographic determinants of infant and early child mortality: a comparative analysis. *Pop Stud* 1985; 39:363-85.
21. Hosmer, D.W. and Lemeshow S. (2000). *Applied Logistic regression.* 2nd Edition*,* John Wiley and Sons, Inc., New York.
22. Hyder AA, Wali SA, McGuckin J: The burden of disease from neonatal mortality: a review of South Asia and Sub-Saharan Africa. Bjog 2003, 110(10):894–901.
23. Ikamari L.D.E., Regional variation in neonatal and post-neonatal mortality in Kenya. *African Population Studies Vol* 27, 1 (Apr 2013).
24. Kamal SMM. Maternal Education as a Determinant of Neonatal Mortality in Bangladesh. DOI: 10.1177/0972063412457509 2012 14: 269 *Journal of Health Management.*
25. Kamal SMM: What is the association between maternal age and neonatal mortality? An analysis of the 2007 Bangladesh Demographic and Health Survey. Asia Pac J Public Health 2011 [http://lib.bioinfo.pl/pmid:22186392].
26. Kousar T, Memon Y, Sheikh S, Memon S, Sehto R (2010). Risk factors and causes of death in Neonates. Department of Pediatrics, Liaquat University of Medical and Health Sciences, Jamshoro.
27. Lawn JE, Cousens S, Zupan J. (2005), 4 million neonatal deaths: when? Where? Why? The *Lancet* 365: 891-900.
28. Mahmood MA. Determinants of neonatal and post-neonatal mortality in Pakistan. *Pak Dev Rev* 2002; 41:723-44.
29. Malqvist M. Neonatal mortality: an invisible and marginalized trauma. *Global Health Action* 2011; 4:5724.
30. Markovitz BP, Cook R, Flick LH, Leet TL. Socioeconomic factors and adolescent pregnancy outcomes: distinctions between neonatal and postneonatal deaths? *BMC Public Health* 2005; 5:79.
31. Mosley WH. and Chen LC. An Analytical Framework for the Study of Child Survival in Developing Countries. *Population and Development Review 1984*; 10: 25–45.

32. Moss W, Darmstadt GL, Marsh DR, Black RE, Santosham M. Research priorities for the reduction of perinatal and neonatal morbidity and mortality in developing country communities. *J Perinatol* 2002; 22:484-95.
33. Oestergaard MZ, Inoue M, Yoshida S, Mahanani WR, Gore FM, Cousens S,Lawn JE, Mathers CD: Neonatal mortality levels for 193 countries in 2009 with trends since1990: a systematic analysis of progress, projections, and priorities. PLoS Med 2011, 8(8):e1001080.
34. Omariba, Walter Rusugu, Fernando Rajulton and Roderic Beaujot, 2008. Correlated mortality of Siblings in Kenya: The Role of State Dependence, *Demographic Research* Vol 18 article 11.
35. Pampel, F. C. 2000. *Logistic Regression: A Primer*. Sage quantitative applications in the Social Sciences Series #132. Thousand Oaks, CA: Sage Publications.
36. Patel TA, Sharma DB. Inter-state variation in Neonatal Mortality Rate among Indian states. Natl J Community Med 2013; 4(1): 54-8.
37. Rahman S, Ansari EL: Neonatal Mortality: Incidence, Correlates and Improvement Strategies. In INTECH. Edited by Oliver CE, Karen OP.; 2012:37–72. ISBN 978-953- 51-0659-3.
38. Rahman S, Salameh K, Bener A, Ansari EL: Socioeconomic associations of improved maternal, neonatal, and perinatal survival in Qatar. Int J Wom Health 2010, 2:311–318.
39. Rajaratnam JK, Marcus JR, Flaxman AD, Wang H, Levin-Rector A, Dwyer L, Costa M, Lopez AD, Murray CJL. Neonatal, post-neonatal, childhood and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. *Lancet* 2010, 375:1988-2008.
40. Regassa N (2012), Infant mortality in the rural Sidama zone, southern Ethiopia: examining the contribution of key pregnancy and postnatal health care services. Journal of Nursing, Social Studies, Public Health and Rehabilitation 1–2, 2012, pp. 51–61.
41. Rutstein S.RO. Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: evidence from the demographic and health surveys. International Journal of Gynecology and Obstetrics (2005) 89, S7-S24.
42. Sathya Susman A (2011). Neonatal and post-neonatal mortality decline in Ethiopia: evidence from DHS 2005. Review of research and social intervention. 32: 44–62.
43. Seedhom AE and Kamal NN (2010). Some determinants of neonatal mortality in a rural area, El-Minia, governorate, Egypt, 2008. The Egyptian Journal of Community Medicine Vol. 28 No.2 April 2010.
44. Singh et al. (2013), Determinants of neonatal mortality in rural India, 2007–2008. PeerJ 1:e75; DOI 10.7717/peerj.75.

45. Susuman, A. S. (2012). Child Mortality Rate in Ethiopia. *Iranian J Publ Health, Vol. 4: 9-19.*
46. Tesfaye Yaekob (2003). Assessment of Pregnancy Outcome with Emphasis on Perinatal and Neonatal Mortality in Dire Dawa Town, Ethiopia. (Unpublished Master Thesis). Addis Ababa University, Ethiopia.
47. Titaley CR, Dibley MJ, Agho K, Roberts CL and Hall1 J. Determinants of neonatal mortality in Indonesia.*BMC Public Health* 2008; 8:232.
48. UNICEF. Committing to Child Survival: A Promise Renewed. Progress Report 2013.
49. UN-IGME (2013). Levels & Trends in Child Mortality Estimates.
50. United Nations. 2001. General Assembly, 56th session. Road map towards the implementation of the United Nations millennium declaration: report of the Secretary-General (UN document no. A/56/326). New York: United Nations.
51. Vittinghoff, E., Shiboski S. C., Glidden D. V., and McCulloch C. E. (2005), Regression Methods in Biostatistics; Linear, Logistic, Survival, and Repeated Measures Models.Springer, New York.
52. Wang PD, Lin RS. Perinatal mortality in Taiwan. *Public Health* 1999; 113:27-33.
53. World Data Bank: World Development Indicators. The World Bank Group; 2013. http://www.worldbank.org/en/country/ethiopia.
54. Zuma N, Lopes C, Kaberuka D, Clark H (2013). Assessing progress in Africa toward the millennium development goals.
55. Zwane E and Masango S (2012). Factors influencing neonatal mortality: an analysis using the Swaziland Demographic and Health Survey 2007. *Journal of Public Health in Africa 2012; 3:e18 doi:10.4081/jphia.2012.e18.*

APPENDIX

Appendix 1

```
proc logistic data=FDG descending;
class Sex(ref="Female") Pre_interval(ref=">36") Birth_Type(ref="Multiple")
Birth_size(ref="Average")  Wealth(ref="Rich") AGE(ref="35-49")
POD(ref="Home") ANC(ref="atleast 4")PEDUC(ref="No education") Bord(ref="5
or more") Region(ref="Addis Ababa");
model AAD= Sex Pre_interval Birth_Type  Birth_size Wealth AGE POD ANC
PEDUC Bord Region/selection=Stepwise lackfit;
Run;
```

Table 1A: Results of multiple logistic regression analysis

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	BIRTH_TYPE		1	1	284.8795		<.0001
2	AGE		2	2	103.2907		<.0001
3	PRE_INTERVAL		2	3	37.6450		<.0001
4	BIRTH_SIZE		2	4	30.6713		<.0001
5	BORD		1	5	16.6585		<.0001
6	ANC		2	6	9.4233		0.0090
7	REGION		10	7	23.3639		0.0095
8	POD		1	8	6.9441		0.0084
9	WEALTH		2	9	8.2589		0.0161

Summary of Stepwise Selection

Step	Variable Label
1	Child is twin
2	Mother's age at child birth
3	Preceding birth interval(months)
4	Size of child at birth
5	Birth order number
6	Antenatal visits during pregnancy
7	Administrative regions
8	Place of delivery
9	Wealth index

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
PRE_INTERVAL	2	45.2869	<.0001
BIRTH_TYPE	1	181.7021	<.0001
BIRTH_SIZE	2	24.6491	<.0001
WEALTH	2	8.2065	0.0165
AGE	2	48.3758	<.0001
POD	1	10.0672	0.0015
ANC	2	9.6909	0.0079
BORD	1	14.9662	0.0001
REGION	10	25.0962	0.0052

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.9968	0.1369	52.9941	<.0001
PRE_INTERVAL	24-36	1	-0.0524	0.0784	0.4468	0.5039
PRE_INTERVAL	<24	1	0.5165	0.0843	37.5178	<.0001
BIRTH_TYPE	Single	1	-1.2972	0.0962	181.7021	<.0001
BIRTH_SIZE	Large	1	0.3824	0.0782	23.9378	<.0001
BIRTH_SIZE	Small	1	-0.1475	0.0844	3.0526	0.0806
WEALTH	Poor	1	0.1974	0.0829	5.6715	0.0172
WEALTH	medium	1	0.0223	0.1055	0.0445	0.8329
AGE	15-19	1	0.9991	0.1592	39.3652	<.0001
AGE	20-34	1	-0.5859	0.0943	38.5827	<.0001
POD	Health facility	1	0.2911	0.0917	10.0672	0.0015
ANC	<4	1	0.0206	0.0881	0.0547	0.8151
ANC	None	1	-0.2457	0.0817	9.0497	0.0026
BORD	1-4	1	0.2801	0.0724	14.9662	0.0001
REGION	Affar	1	-0.3379	0.2192	2.3753	0.1233
REGION	Amhara	1	0.5083	0.1543	10.8497	0.0010
REGION	Benishangul-Gumuz	1	0.4820	0.1664	8.3931	0.0038
REGION	Dire Dawa	1	-0.7486	0.3048	6.0338	0.0140
REGION	Gambela	1	0.0396	0.1919	0.0426	0.8365
REGION	Harari	1	0.1383	0.2182	0.4017	0.5262
REGION	Oromiya	1	0.2020	0.1466	1.8974	0.1684
REGION	SNNP	1	0.1591	0.1572	1.0234	0.3117
REGION	Somali	1	0.0797	0.1937	0.1692	0.6808
REGION	Tigray	1	0.1072	0.1712	0.3919	0.5313

Odds Ratio Estimates

Effect			Point Estimate	95% Wald Confidence Limits	
PRE_INTERVAL	24-36 vs >36		1.509	1.157	1.969
PRE_INTERVAL	<24 vs >36		2.666	2.003	3.549
BIRTH_TYPE	Single vs Multiple		0.075	0.051	0.109
BIRTH_SIZE	Large vs Average		1.854	1.423	2.415
BIRTH_SIZE	Small vs Average		1.091	0.820	1.452
WEALTH	Poor vs Rich		1.517	1.139	2.021
WEALTH	medium vs Rich		1.274	0.887	1.829
AGE	15-19 vs 35-49		4.106	2.412	6.987
AGE	20-34 vs 35-49		0.841	0.620	1.141
POD	Health facility vs Home		1.790	1.249	2.564
ANC	<4 vs atleast 4		0.815	0.586	1.133
ANC	None vs atleast 4		0.624	0.458	0.851
BORD	1-4 vs 5 or more		1.751	1.318	2.326
REGION	Affar	vs Addis Ababa	1.339	0.517	3.466
REGION	Amhara	vs Addis Ababa	3.120	1.288	7.558
REGION	Benishangul-Gumuz	vs Addis Ababa	3.039	1.242	7.438
REGION	Dire Dawa	vs Addis Ababa	0.888	0.319	2.469
REGION	Gambela	vs Addis Ababa	1.953	0.785	4.858
REGION	Harari	vs Addis Ababa	2.155	0.864	5.378
REGION	Oromiya	vs Addis Ababa	2.297	0.956	5.521
REGION	SNNP	vs Addis Ababa	2.200	0.903	5.361
REGION	Somali	vs Addis Ababa	2.032	0.808	5.115
REGION	Tigray	vs Addis Ababa	2.089	0.855	5.104

Association of Predicted Probabilities and Observed Responses

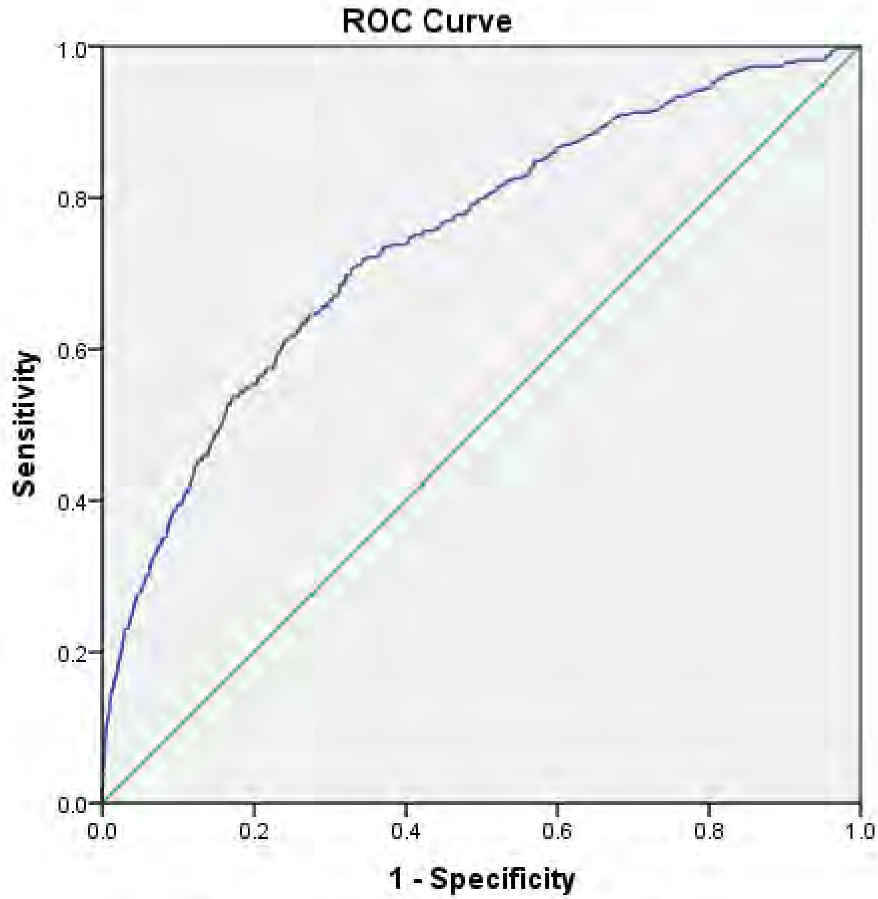
Percent Concordant	73.7	Somers' D	0.488
Percent Discordant	24.9	Gamma	0.495
Percent Tied	1.4	Tau-a	0.055
Pairs	2239650	c	0.744

Partition for the Hosmer and Lemeshow Test

		AAD = Yes		AAD = No	
Group	Total	Observed	Expected	Observed	Expected
1	631	10	9.47	621	621.53
2	631	12	13.73	619	617.27
3	632	12	17.23	620	614.77
4	628	22	19.83	606	608.17
5	625	26	23.44	599	601.56
6	633	18	28.02	615	604.98
7	629	34	32.81	595	596.19
8	633	41	41.02	592	591.98
9	630	70	53.77	560	576.23
10	631	133	138.70	498	492.30

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
11.8764	8	0.1568



Diagonal segments are produced by ties.

Table 1B: Results of diagnostic tests for outliers and influential values

Descriptive Statistics

	N	Minimum	Maximum
Analog of Cook's influence statistics	6303	.00000	.18996
Leverage value	6303	.00071	.05188
DFBETA for constant	6303	-.09136	.16059
DFBETA for Pre_interval(1)	6303	-.01375	.01452
DFBETA for Pre_interval(2)	6303	-.01236	.01192
DFBETA for Birth_Type(1)	6303	-.02680	.02261
DFBETA for Birth_size(1)	6303	-.01119	.01393
DFBETA for Birth_size(2)	6303	-.00980	.01192
DFBETA for Wealth(1)	6303	-.01888	.01541
DFBETA for Wealth(2)	6303	-.02194	.03101
DFBETA for AGE(1)	6303	-.03308	.04470
DFBETA for AGE(2)	6303	-.02139	.01656
DFBETA for POD(1)	6303	-.03410	.03121
DFBETA for ANC(1)	6303	-.02020	.02218
DFBETA for ANC(2)	6303	-.01731	.01695
DFBETA for Bord(1)	6303	-.01865	.01685
DFBETA for Region(1)	6303	-.16767	.07957
DFBETA for Region(2)	6303	-.16633	.08036
DFBETA for Region(3)	6303	-.16795	.07704
DFBETA for Region(4)	6303	-.16720	.07807
DFBETA for Region(5)	6303	-.16469	.07989
DFBETA for Region(6)	6303	-.16699	.07898
DFBETA for Region(7)	6303	-.16820	.07845
DFBETA for Region(8)	6303	-.16392	.08018
DFBETA for Region(9)	6303	-.16148	.07996
DFBETA for Region(10)	6303	-.16004	.11662
Valid N (listwise)	6303		

Declarations

I, the undersigned, declare that the thesis is my original work, has not been presented for a degree in any other university and that all sources of material used for the thesis have been duly acknowledged.

Name: Garoma Wakjira

Signature: ______________________

Date: June, 2014

This thesis has been submitted for examination with my approval as a University advisor.

Advisor: Prof M.K Sharma

Signature: _________________________

Date: June, 2014