Modeling Neonatal Mortality in a Teaching Hospital in Ghana

Benjamin Atta Owusu

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Applied Mathematics

Prince of Songkla University

2016

Thesis Title Modeling Neonatal Mortality in a Teaching Hospital in Ghana

Author Mr. Benjamin Atta Owusu

Major Program Applied Mathematics

Major Advisor:	**Examining Committee:**
……..	……Chairperson
(Asst. Prof. Dr. Apiradee Lim)	(Asst. Prof. Dr. Aniruth Phon-On)
	……
Co-advisor:	(Asst. Prof. Dr. Apiradee Lim)
……..	……
(Dr. Nifatamah Makaje)	(Dr. Nifatamah Makaje)
	……
Co-advisor:	(Asst. Prof. Dr. Areeyuth Sama-Ae)
……..	……
(Asst. Prof. Dr. Areeyuth Sama-Ae)	(Asst. Prof. Dr. Bungon Kumphol)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Applied Mathematics.

……

(Assoc. Prof. Dr. Teerapol Srichana)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

……Signature
(Asst. Prof. Dr. Apiradee Lim)
Major Advisor

……Signature
(Benjamin Atta Owusu)
Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

…….……Signature

(Benjamin Atta Owusu)

Candidate

ชื่อวิทยานิพนธ์ แบบจำลองการตายของทารกแรกเกิดในโรงพยาบาลฝึกหัดแพทย์ในประเทศกานา

ผู้เขียน Benjamin Atta Owusu

สาขาวิชา คณิตศาสตร์ประยุกต์

ปีการศึกษา 2559

บทคัดย่อ

ยี่สิบแปดวันแรกของทารกแรกเกิด คือช่วงเวลาที่เสี่ยงมากที่สุดสำหรับการรอดชีวิตของทารก อัตราการตายของทารกแรกเกิดมีแนวโน้มลดลงเมื่อไม่กี่ปีที่ผ่านมา สิ่งสำคัญคือความเข้าใจในปัจจัยเสี่ยงที่มีความสัมพันธ์กับอัตราการตายของทารกแรกเกิด ณ หน่วยบริบาลทารกแรกเกิด เพราะแสดงถึงคุณภาพของการดูแลทารกแรกเกิด วัตถุประสงค์หลักของการศึกษาในครั้งนี้ คือ เพื่อระบุอัตราการตายของทารกแรกเกิดและค้นหาปัจจัยเสี่ยงที่ส่งผลต่อการอยู่รอดของทารก ในระดับตติยภูมิสถานพยาบาลที่เข้ารับการรักษา ด้วยการนำตัวแบบทางคณิตศาสตร์มาใช้ในการอธิบายความสัมพันธ์ของปัจจัยดังกล่าว ข้อมูลทุติยภูมิได้รับจากหน่วยบริบาลทารกแรกเกิด ในช่วงเวลา 17 เดือน จากมกราคม ปี 2556 ถึง พฤษภาคม ปี 2557 จำแนกข้อมูลออกเป็น 2 กลุ่ม ตามกับสถานที่คลอดของทารก สร้างตัวแบบการถดถอยลอจิสติกสำหรับข้อมูลทั้ง 2 กลุ่ม เพื่อหาความสัมพันธ์ระหว่างการตายของทารกแรกเกิดและปัจจัยเสี่ยง และสร้างตัวแบบการถดถอยลอจิสติกสำหรับข้อมูลทั้งหมด เพื่อหาปัจจัยเสี่ยงต่อการตายของทารกแรกเกิด ณ หน่วยบริบาลทารกแรกเกิด ตรวจสอบฟังก์ชันความน่าจะเป็น และฟังก์ชันความควรจะเป็นที่สัมพันธ์กับตัวแบบการถดถอยลอจิสติก

ใช้กระบวนการ Fisher's scoring เพื่อหาค่าเหมาะสมที่สุดในการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยลอจิสติก และวิธี Newton Raphson's ด้วยขั้นตอนวิธีการของ Marquardt damping ในการประมาณค่าสัดส่วนที่ปรับค่าของตัวแปร ตัวแบบลอจีสติกแต่ละตัวแบบตรวจสอบข้อตกลงของการแจกแจงปกติของค่าความคลาดเคลื่อนด้วยการสร้าง Normal quantile-quantile plot และตรวจสอบความกลมกลืนและการทำนายของตัวแบบด้วยการสร้าง Receiver Operating Characteristics plots

การวิเคราะห์เชิงเส้นกำกับสำหรับฟังก์ชันความควรจะเป็นสำหรับตัวแบบการถดถอยลอจิสติกแสดงให้เห็นว่าเป็นไปตามข้อตกลงของตัวแบบในทุกเงื่อนไข ที่ได้จากการประมาณค่าเดียวของค่าความควรจะเป็นสูงสุด การประมาณค่าที่ได้ถือว่าเป็นเชิงเส้นกำกับปกติ และมีความสอดคล้องสำหรับการอธิบายการตายของทารกในกลุ่มข้อมูลที่แตกต่างกัน น้ำหนักทารกแรกคลอด คะแนนการประเมินสภาวะทารกแรกเกิดใน 5 นาที และผลการวินิจฉัยก่อนออกจากหน่วยสถานพยาบาล พบว่าเป็นปัจจัยเสี่ยงที่มีความสัมพันธ์กับการตายของทารกที่คลอดในสถานพยาบาลต่าง ๆ แล้วถูกส่งต่อมาที่หน่วยบริบาลทารกแรกเกิด ตัวแปรดังกล่าวข้างต้นทั้ง 3 ตัวแปร และอายุครรภ์มีความสัมพันธ์อย่างมีนัยสำคัญกับการตายของทารกที่ไม่ได้ส่งต่อจากสถานพยาบาลอื่น สถานที่คลอดของทารกแรกเกิดเป็นปัจจัยหนึ่งที่มีผลอย่างมีนัยสำคัญต่อโอกาสการรอดของทารกในหน่วยบริบาลทารกแรกเกิด อัตราการตายทารกแรกเกิดในภาพรวม ณ หน่วยบริบาลทารกแรกเกิดยังคงสูงมาก จึงจำเป็นที่จะต้องคำนึงและให้การช่วยเหลือเพื่อลดภาวะเสี่ยงที่จะเกิดขึ้นกับทารกแรกเกิด

Thesis Title Modeling Neonatal Mortality in a Teaching Hospital in Ghana

Author Benjamin Atta Owusu

Major Program Applied Mathematics

Academic Year 2016

ABSTRACT

The first 28 days of life- the neonatal period is the most vulnerable time for a child's survival. Neonatal mortality has seen a downward trend in recent years. Understanding the risk factors associated with neonatal mortality at the neonatal unit is important because it allows inferences about the quality of care. The main objective of this study was to determine the neonatal mortality rate and to provide information on the risk factors that affects the survival of neonates at the neonatal unit of a tertiary health facility and to present a mathematical model for explaining such factors. Secondary data were obtained from the neonatal unit of a hospital over a seventeen-month period, from January 2013 through to May 2014. The data was classified into two, based on the place of delivery of these neonates. Logistic regression models were fitted to the two datasets to assess the association between neonatal death and its risk factors. Another logistic regression model was fitted to the combined dataset in order to investigate the general risk factors of neonatal death at the neonatal unit. The probability density functions and likelihood functions associated with the logistics regression models were verified for regularity. Fisher's scoring optimization algorithm was used to estimate the parameters of the logistic

regression models and Newton Raphson's method with Marquardt damping algorithm was used to estimate adjusted proportions of the variables. Each logistic regression model was further assessed for asymptotic normality using the normal quantile-quantile plots. Using Receiver Operating Characteristics plots, the predictive power of the models were also estimated.

Asymptotic analysis of the likelihood function for the logistic regression model, showed that, the function satisfied all the regularity conditions for the existence and uniqueness of maximum likelihood estimates. The estimates were also found to be asymptotically normal and consistent for interpreting the neonatal mortality for the different datasets.

Birth weight, 5 minute Apgar score and discharge diagnosis were found to be the main risk factors associated with neonatal death among babies who were delivered at different health facilities and referred to the unit. These three variables together with gestational age were significantly associated with mortality of babies who were not referred from different facilities. The place of delivered of the neonates also had significant effect on their survival chances at the neonatal unit. The overall mortality rate at the neonatal unit is very high. There is the need for urgent attention and interventions to help reduced the risk associated with these neonates.

Acknowledgement

Foremost, I thank the Almighty God for his Grace and Mercies on my life. I would like to express my sincere gratitude and appreciation to my advisors, Assist. Prof. Dr. Apiradee Lim, Assist. Prof. Dr. Areeyuth Sama-Ae and Dr. Nifatamah Makaje for the continuous and consistent support for my MSc study and research. Their patience, motivation, enthusiasm, difficult questions, guidance and immense knowledge helped me in every stage of research and writing of this thesis. Their continuous encouragement and challenge was the bedrock of my progress throughout the preparation of this work. I could not have imagined having a better advisors and mentors for my MSc study.

Secondly, I would like to thank the Applied Mathematics program committee chaired by Assist. Prof. Dr. Arthit Intarasit for affording me the opportunity to broaden my knowledge in this program. My gratitude, again, goes to the Department of Mathematics and Computer Science through Assist. Prof. Dr. Areeyuth Sama-Ae (Head of Department) for providing the serene environment for my study and research. I appreciate the opportunity I had to tap into the faculty especially, Sunaree Suwanro; she became a big sister to me. My sincere thanks also goes to Emeritus Prof. Donald McNeil (The Don) for assisting me with his deep knowledge in research methods. Next, I would like acknowledge the Higher Education Research Promotion and the Thailand's Education Hub for Southern Region of ASEAN Countries Project Office of the higher Education Commission for providing funds for this research. I am thankful to the Faculty of Science and Technology and Office of the Graduate School, Pattani campus.

I thank my colleague course mates in Applied Mathematics and Research Methodology programs for the sleepless nights we were working together at nothing to meet deadlines, for all the fun we had, the academic setbacks and disappointments we shared in the last two academic years.

Finally, I would like to say a big thank you to everyone who made the slightest effort to assist my work.

Benjamin Atta Owusu

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Neonatal mortality is the death occurring among children who are less than 4 weeks old (Lander, 2006). The first four weeks of life (neonatal period) is the most crucial period when infants are highly prone to illness and death. Latest reports indicate that over 5.9 million under-five deaths were recorded in 2015 with daily average of 16,000 deaths (UNICEF, 2015). Globally, almost 3 million babies die in the neonatal period each year (Shiffman, 2015). This accounts for almost 44% of under-five deaths (Kinney *et al*., 2015).

Majority of these neonatal deaths occur in developing countries. In fact, neonatal mortality accounts for almost half of all infant deaths in these countries (Klingenberg *et al.*, 2003). Preventive measures and interventions have helped to reduce the global neonatal mortality rate from 36 in 1990 to 19 per 1,000 live births in 2015 (You *et al*., 2015). Globally, neonatal mortality has been decreasing at a slower rate from 1990 as compared to under-five mortality (Lawn *et al*., 2014). The decline in neonatal death especially in Ghana from 32 per 1,000 live births in 2008 to 29 per 1,000 live births in 2014 has partly been attributed to improved facilities at the neonatal units and improved neonatal care (Enweronu-Laryea *et al*., 2008; GDHS, 2009; GDHS, 2014).

However, neonatal mortality still remains a major challenge for Ghana and other lower-middle income countries particularly in Sub-Saharan Africa and southern Asia;

countries in these two regions have made the least progress towards reducing neonatal mortality (Dickson *et al*., 2014). Less than 100 countries were able to achieved Millennium Development Goal-4 (MDG 4) of reducing under-five mortality by two-thirds (UNICEF, 2015). Through the newly formulated Sustainable Development Goals (SDG), United Nations (UN) member nations have committed themselves to reducing under-five mortality and neonatal mortality to 25 and 15 deaths per 1,000 live births respectively by the year 2030 (Taylor *et al*., 2015). Meeting this new target is much dependent on understanding and reducing the risk of death among newborns at the various neonatal units. Infant and neonatal mortality rates are significant measures of health quality, societal welfare and socio-economic status of a country (Arafa and Alshehri, 2003; Hsu *et al*., 2014; Kazembe and Kandala, 2015). Therefore, attention must be given to the infant who is less than 28 days (neonate).

The survival of a child beyond five years depends much on his survival during the neonatal period. Examining the risk factors of neonatal mortality at the various neonatal units is necessary as it allows inferences about the quality of care. It could provide insights into how neonates could be managed to improve the outcomes of admissions at the neonatal unit. Although a lot of studies have been conducted on the subject of neonatal mortality at the neonatal intensive care unit (NICU), almost all these studies estimate neonatal mortality rate based on number of live births. However, the neonatal unit is not and has never been a place of delivery and, therefore, estimating mortality rate at the unit based on live births gives room for incorrect reporting and interpretation of the real problem. Separate studies have identified place of birth, birth weight, mode of delivery, delayed breastfeeding

initiation and age of baby at the time of admission as significant predictors of neonatal death (Ajaari *et al.*, 2012; Bacak *et al.,* 2005; Edmond, 2006; Hsu *et al.,* 2014).

The consistent evolution statistical and mathematical models for interpreting the causes and risk factors diseases and death means that statisticians and mathematicians need to develop new and improved methods for neonatal mortality. Data on admission outcomes at the neonatal unit are considered dichotomous, and therefore, any models that seek to explain the risk of death much be accurate, sensitive and specific. A lot of models have been developed to understand the risk of mortality in the neonatal unit. Logistic regression analysis has become the most efficient tool in analysis data with dichotomous (binary) outcomes. A lot of literature has been written on the development of the model and it's uses. Just like any generalized linear model, the logistic regression model has some very important underlying assumptions that ensures that the estimates from the model are accurate. However, most researchers who have used the logistic regression model pay less attention to the regularity conditions of the functions involved in estimating the parameters and asymptotic properties of the estimates from the model. A standard procedure has been developed for estimating parameters of various regression models. This procedure is the Maximum Likelihood estimation method and it has become the best approach to finding the parameters of the logistic regression model (van der Vaart, 1998). This thesis uses the logistic regression model to find the factors associated with neonatal mortality and applies the method of maximum likelihood estimations to find and interpret the parameter estimates from the model.

The aim of this dissertation is to investigate the major causes and risk factors of neonatal mortality and estimate the neonatal mortality rate based on the number of babies admitted to the Mother and Baby Unit (MBU), Komfo Anokye Teaching Hospital (KATH), Kumasi, Ghana. This dissertation also provides analytical inferences of the logistic regression model. Theoretical proves of the regularity conditions of the likelihood function for the logistic regression model are also presented in this dissertation.

1.2 Objectives

The main objectives of the study are:

i. To present the method of estimating the parameters of logistic regression.
ii. To examine the neonatal mortality rate at the neonatal unit of a teaching hospital in Ghana.
iii. To identify the factors associated with neonatal mortality in a teaching hospital in Ghana.

1.3 Literature Review

Different researchers have taken time to examine neonatal mortality, the predictors, the determinants and their risk factors using health survey data. Neonatal mortality is the death occurring among children who are less than 4 weeks old (Lander, 2006). It could also be defined as death within the first 27 days of birth (Goodman *et al*., 2002).

1.3.1 Neonatal mortality trends

The close of 2015 brought down the curtains on MDGs, and policy makers accessed the impacts and progress made by MDG-4 on child health and made formulations for the post 2015 agenda (UNICEF, 2015). Neonatal mortality has seen a slower rate of reduction than both maternal and under-five mortality. Highly burden countries, especially in Africa have seen the least progress in reducing neonatal deaths. These deaths happen at a fast rate and a quick response from health workers and health care providers to reduce it. The major causes of the 2.9 annual neonatal deaths have been infections, problems of preterm birth and intrapartum conditions (Lawn *et al.*, 2014).

Providing effective care for women who deliver in health care facilities and their newborns could save over 1.3 million neonates annually by the year 2020. Government and health policy makers should extend the coverage and quality of antenatal and postnatal interventions as this can avert 71% of all neonatal deaths before 2025 (Bhutta *et al.*, 2014). The decline rate of neonatal mortality in Africa has been the slowest compared with other regions of the world. Africa accounts for over 30% of global neonatal mortality. At the current pace of reduction in neonatal mortality, it will take over hundred years for a newly born African baby to have the same odds of survival with a baby delivered in Europe or North America (Lawn *et al.*, 2014).

1.3.2 Neonatal factors

Neonatal and perinatal care has been regionalized effectively in Canada with at least 17 neonatal intensive care units across the country. In 2002, Sankaran *et al*., (2002)

used a 22 months' retrospective data (January 1996 to October 1997) from all the 17 neonatal intensive care units to determine the differences in neonatal mortality rates. In the process, they used information on 19,265 infants admitted to these units. They employed univariate and bivariate analysis to determine the characteristics of the study sample and to measure the level of correlation between these characteristics and neonatal death. Their study revealed that the overall mortality rate among the 17 units was 4 per 1,000 live births and the most prevalent neonatal conditions associated with neonatal death were gestational age less than 24 weeks, out-born status, and congenital anomalies. Their analysis showed that Apgar score less than 7 is also significantly associated with neonatal death. They also found that 40% of deaths occurred within the first 48 hours of admission. Therefore, they recommended that a lot of attention should be given to preterm babies to increase their chances of survival (Sankaran *et al*., 2002).Adetola *et al*, conducted a study in 2011 to determine the neonatal mortality rate, causes of death and the associated risk factors among hospital live births in Ibadan, Nigeria. Their findings revealed that, out of 1,058 live births, preterm, low birth weight, severe perinatal asphyxia and infections are some of the leading neonatal causes and risk factors of death. They concluded that there is a need for programs to improve the skills of neonatal resuscitation and care for LBW infants (Adetola *et al.*, 2011).

In 2014, Kayode *et al*, decided to identify individual and community determinants of neonatal mortality. They revealed that individual neonatal characteristics such as place of delivery, 5th minute Apgar score, birth weight and gestation have much effect on neonatal death (Kayode *et al*., 2014). Neonatal mortality accounts for more than half of all

under-five deaths in developing countries mostly in sub-Saharan Africa. In a three-month study aimed at describing the mortality and morbidity patterns in a Tanzanian special care baby unit, almost 250 neonatal admissions were audited. The study site was Kilimanjaro Christian Medical Centre. The study revealed that low birth weight, prematurity and infections account for over 60% of neonatal admissions to the unit (Klingenberg *et al*., 2003).

1.3.3 Maternal factors

Mothers with birth complications have significantly high risk of neonatal mortality in Pakistan. It is very important to implement interventions that focus on antenatal care, and managing birth-associated problems in order to reduce neonatal mortality. Effective referral systems can also help in reducing neonatal death. Nisar and Dibley (2014) used data from the Pakistan Demographic and Health Survey 2006-07 to find the determinants of neonatal mortality in Pakistan. They concluded that maternal factors increase the risk of neonatal mortality (Nisar and Dibley, 2014).

With combined data from the Ghana Demographic and Health Survey (GDHS) 2003 and 2008, hierarchical modeling was used to examine 6900 women. Kayode *et al*., (2014), concluded that using antenatal, delivery and effective postnatal health services reduced the risk of neonatal death (Kayode *et al.*, 2014).

1.3.4 Methods

Czepiel (2002) investigated the reasons why logistic regression models are mostly used to model categorical variables with dichotomous responses. He found that the

response values of categorical data are not measured on a ratio scale and their errors too are not normally distributed. He also noted that the theory of generalized linear models has identified vital properties that are common to a class of probability distributions which includes the binomial model. He also presented a mathematical process to estimate the parameters of the binomial model which included finding maximum likelihood estimates of the log-likelihood function for logistic regression and applying iterative methods. (Czepiel, 2002).

Mai et al (2014) presented optimization algorithms for maximum likelihood estimations. They formulated the log-likelihood function of any generalized linear model as a system of unconstrained optimization problem and presented methods for solving such problems. Their research presented linear search algorithms, gradient algorithms and Hessian approximation algorithms for solving maximum likelihood problem (Mai *et al,* 2014)

In a study to estimate the risk factors of gestational diabetic mellitus, Okeh and Oyeka (2013) presented a matrix approach to solving the unconstrained log-likelihood function to estimate the parameters of logistic regression. They used Newton's method and a Quasi-Newton method to solve the log-likelihood function. Their study concluded that the Quasi-Newton approach assures convergence of the iterative process (Okeh and Oyeka, 2013).

1.4 Scope of the study

The analysis will focus on 5,195 babies admitted to the Mother and Baby Unit (MBU) of Komfo Anokye Teaching Hospital (KATH). The determinants of the study are birth weight, delivery mode, gestational age, place of delivery and admission age. Retrospective primary data collected from the unit from January 2013 to May 2014 will be used in this study. The strength of association between the determinants and neonatal mortality was accessed with logistic regression. After this, maximum likelihood method was used to approximate the coefficients of each determinant in the model while using Fisher's scoring iteration to approximate the coefficients for all the data set. Newton-Raphson's iteration with Marquardt damping algorithm was used to find adjusted proportions for each variable.

1.5 Outline of Chapters

This dissertation is presented in the following structure: Chapter 1 contains the background and motivation of this study. It explains the problem of neonatal mortality in the neonatal unit, objectives of the study and presents literature that is relevant to the topic. Chapter 2 provides detailed methods formulating and explaining the predictive logistic regression model. In addition, it provides detailed methods of estimating the parameters of the logistic regression model and an overview of the asymptotic properties of the parameter estimates from the model. Chapter 3 contains information of the asymptotic properties of the logistic regression model. The results after applying the model to real life data is also

presented in this chapter. Chapter 4 contains discussions of the results, impact on public health administration and conclusions and recommendations of the study.

Chapter 2

Methodology and Basic Knowledge

2.1 Population and Sample

This is a hospital based retrospective study of 5,195 neonates admitted to the neonatal unit of a teaching hospital. Data were collected from the neonatal unit of Komfo Anokye Teaching Hospital (KATH), Kumasi Ghana. The KATH is a tertiary health facility located in a Kumasi, the capital of the Ashanti region. Kumasi is also the second most populous city in Ghana. KATH serves as the main referral destination for other hospitals in the region as well as the northern sector of Ghana. It has an equipped neonatal unit, also called the Mother and Baby Unit (MBU). The neonatal unit has three wards. There is a High Dependency Unit (HDU) that admits sick babies referred from the delivery wards. There is also a Preterm/Low birth weight (LBW)/Kangaroo Mother Care (KMC) unit that admits preterm/LBW babies from the HDU who have been stabilized and there is a septic unit that admits out-born sick newborns and infants up to the age of two months. Different categories of babies are admitted to the facility. These include preterm babies, low birth weight babies, babies with neonatal jaundice and risk of sepsis, babies with congenital anomalies.

2.1.1 Approach to data collection and sample

All neonates admitted to the unit during the period from January 2013 through to May 2014 were included in the study. They were followed until discharged or death. The following information were retrieved from the in-patient files: birth weight, sex, mode of

delivery, place of delivery, gestational age, age at admission, Apgar scores and outcome of admission. Neonatal mortality is defined as death that occurred within the first 28 days of life. This study targets only babies who 28 days or less, hence all babies who are more than 28 days are excluded from the study. The data was then divided into two, based on place of delivered. Place of delivery is categorized into two: either delivered at the KATH or is referred from a different health facility. A model is fitted to both datasets and the results compared. The overall management of data is presented as a flow diagram in Figure 2.1.

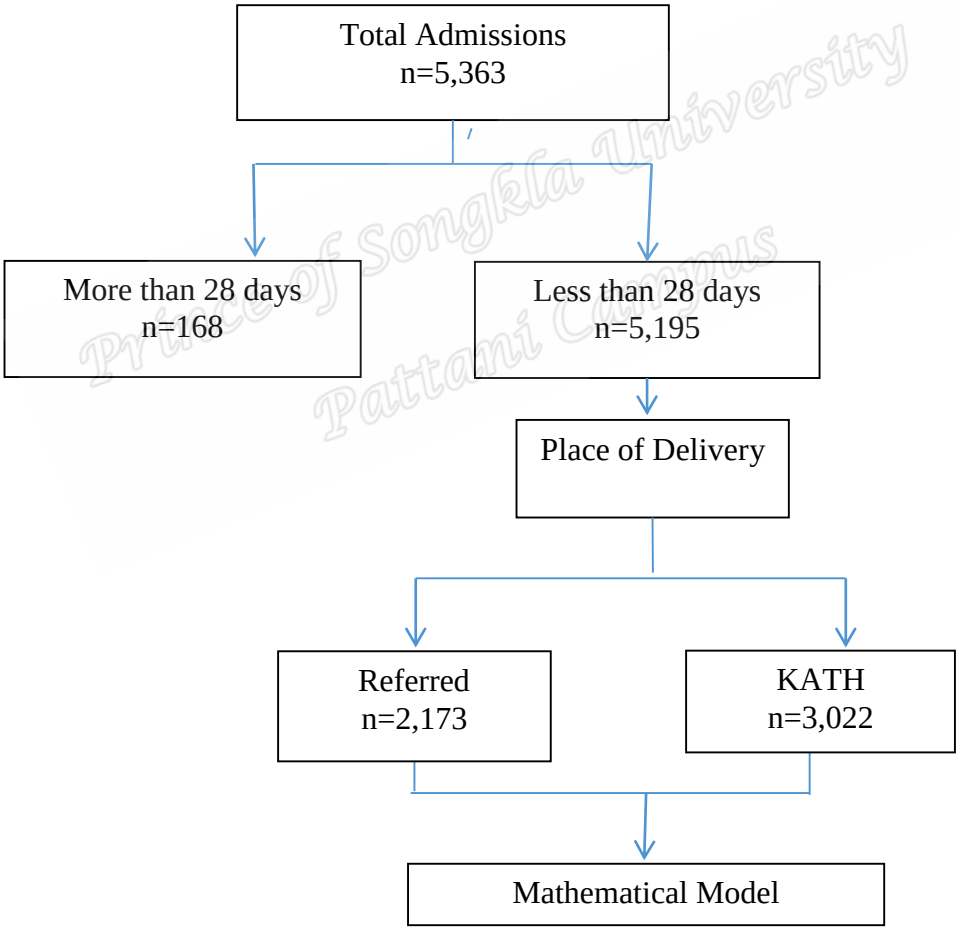

Figure 2.1 Flow diagram of data management

During the period, there were 5,363 babies admitted to the unit. The overall management of data is presented as a flow diagram in Figure 2.1. The birth weight is categorized into four groups namely; very low birth weight (VLBW), low birth weight (LBW), and normal weight. VLBW is defined as weights less than 1.5 kg; LBW for those between 1.5-2.4 kg, normal birth weight for those above 2.4 kg. The mode of delivery is divided into three categories. The first group is spontaneous vaginal delivery (SVD) which includes all neonates delivered by spontaneous delivery. The next category is vacuum. This is made up of all babies delivered by vacuum extraction. Babies delivered through caesarean section (C/S) are also categorized as a single group. Apgar score at 5 minutes is categorized in three groups; 0-3, 4-7 and above 7. Babies' gestational age is classified as preterm (< 36 weeks), term (> 36weeks) or not stated. The discharge diagnosis is also classified into eight groups. The most prevalent conditions were prematurity, respiratory distress, infections, congenital anomalies, neonatal jaundice and birth asphyxia. The age as of the time of admission was divided into four groups; 1 day, 2-7 days, 8-14 days and 15-28 days.

2.1.2 Path diagram for the variables

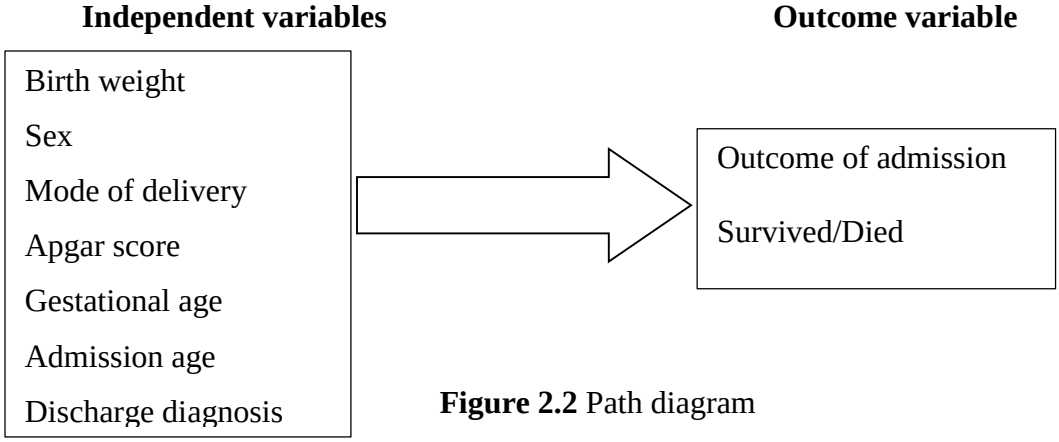

Figure 2.2 Path diagram

2.2 Ethical approval

This study was approved by the Research and Development Unit of KATH and the Committee on Human Research, Publications and Ethics, Kwame Nkrumah University of Science and Technology, School of Medical Sciences and KATH, Kumasi. The research protocol was reviewed and accepted on 10th November, 2015 with reference number CHRPE/AP/365/15.

2.3 Data Analysis

2.3.1 Generalized linear models

Every analytic procedure that uses regression models is preceded by a comparison of each variable with the outcome of interest. This is also univariate analysis. The univariate analysis helps to select the factors that will be used in the multivariate analysis (Abreu *et al*., 2009). The chi-square test is one of the univariate analysis methods used to select factors since it considers the ordinal natural of the response variable. Normally, a conservative level of significance is used (generally between 10% and 25%) for entering the factors (variables) in the model (Hosmer *et al*., 2004). The best way to explain neonatal mortality is by using a linear relationship between the determinants and outcome. However, the understanding such a linear model is not very easy when considering different data frames and risk factors. More often, there are interactions between these risk factors and this hinders the authenticity of the results. Hence, the logistic regression model which is part of the family of generalized linear models is widely used to analyze data relating to epidemiological studies.

The equations used to analyze epidemiological data are forms of generalized linear models. A generalized linear model consists of three components:

1. A random component: this specifies the conditional distribution of a response variable, Y_i subject to some explanatory variables.
2. A linear predictor which is a function of the variables. The linear predictor is usually of the form:

   $$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_j x_j,$$

   where x_j are explanatory variables.
3. A smooth link function g. The link function is also invertible. This function transforms the expectation of the response variable $\mu_i = E(Y_i)$ to the linear predictors:

   $$g(\mu_i) = y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_j x_j.$$

The link function is invertible and hence its inverse exists and can be written as:

$$\mu_i = g^{-1}(y_i) = g^{-1}(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_j x_j).$$

Table 2.1 Some link functions and their inverses.

Link	$y_i = g(\mu_i)$	**Inverse**: $\mu_i = g^{-1}(y_i)$
Identity	μ_i	y_i
Log	$\ln(\mu_i)$	$\exp(y_i)$
Inverse	μ_i^{-1}	y_i^{-1}
Inverse-square	μ_i^{-2}	$(y_i)^{-1/2}$
Square-root	$\sqrt{\mu_i}$	y_i^2
Logit	$\ln\left(\frac{\mu_i}{1-\mu_i}\right)$	$\frac{1}{1+\exp(y_i)}$
Log-log	$-\ln[-\ln(\mu_i)]$	$\exp[-\exp(y_i)]$
Complementary log-log	$\ln[-\ln(1-\mu_i)]$	$1-\exp[-\exp(y_i)]$
Probit	$\varphi^{-1}(\mu_i)$	$\varphi(y_i)$

Source: Fox, (2008)

2.3.2 The logistic regression model

Logistic regression is used to model dichotomous or binary outcome with takes values 0 and 1(Bewick *et al.*, 2005). It studies the relation between a dependent variable (dichotomous) and a group of independent variables. The responses of linear models are not measured on a ratio scale and their errors are not normally distributed (Czepiel, 2002). Hence, linear models are inappropriate for the modeling of categorical variables. The logistic regression is derived from the binomial probability distribution and it is part of the family of generalized linear models (GLMs). It uses logit transform (natural logarithm of odds) to show that a particular event will occur.

2.3.2.1 Model formulation

Suppose that the observations under study can be categorized according to the variables of interest and then aggregated into b groups in such a way that all observations in a group have identical values for all the variables of interest. Let n_i be the number of observations (babies) in group i such that $\sum_{i=1}^{b} n_i = N$ (sample size) and let y_i be the number of observations who have the attribute of interest (died) in group i. Thus y_i is the number of babies who died in group i. The y_i is a realization of a random variable Y_i that takes the values $0,1,2,...,n_i$. The n_i observations in each category are independent and have the same observed probability p_i of death. Hence, the probability distribution function of Y_i is binomial and given by

$$\Pr\{Y_i = y_i\} = \binom{n_i}{y_i} p_i^{y_i} (1-p_i)^{n_i - y_i}, \qquad (2.1)$$

for $0 \le y_i \le n_i$. Here $p_i^{y_i}(1-p_i)^{n_i-y_i}$ is the probability of obtaining y_i successes and $n_i - y_i$ failures in some specific order.

Let $\mathbf{y}$ be a column vector containing elements y_i representing the observed counts of the number of successes for each group. The logistic regression model which is formulated from the probability distribution function has a linear component which contains a design matrix and a vector whose parameters will be approximated. The logistic regression model establishes a relation between the logit transformation, the log-odds of the probability of success and the linear component.

Thus it moves from probability p_i to odds ratio OR; i.e. $OR_i = \frac{p_i}{1-p_i}$. The log odds of success is given by

$$\ln\left(\frac{p_i}{1-p_i}\right) = \sum_{j=0}^{K} x_{ij}\beta_j \quad \text{for } i = 1,2,...,b \tag{2.2}$$

where K is the number of independent variables of interest.

Let β be a column vector of parameters β_j. Each of the K independent variables has one parameter that corresponds to it and β_0 for the intercept. The parameters $\beta_0, \beta_1, \beta_2, ..., \beta_K$ are to be determined. Equation (2.2) is best applied when considering continuous dependent variables. However, for categorical variables, dummy variables are created based on the number of categories in each variable. A dummy variable is any variable that takes on a finite number of values so that different categories of a variable can be identified (Kleinbaum, 1998). For an independent variable with A categories, then exactly $A-1$ dummy variables must be defined to index these categories. Thus, if there is one independent variable with 4 categories, then 3 dummy variables will be included in the regression model. This is because; one category is removed from the regression model to serve as a reference group.

Let $\mathbf{X}$ be a design matrix with element x_{ij}. The first column elements of each row in the design matrix $x_{i0} = 1$. And $\mathbf{X}$ has dimensions b rows and m columns, where $m = \sum A - K$ is the number of dummy variables and the column vector β of parameters has the dimensions m.

Let $\mathbf{p}$ be a column vector of length b with elements $\ln\left(\frac{p_i}{1-p_i}\right)$. Therefore, the logistic regression model is presented as

$$\ln\left(\frac{p_i}{1-p_i}\right)=\sum_{j=0}^{m}x_{ij}\beta_j \quad \text{for } i=1,2,...,b \tag{2.3}$$

or in matrix form as

$$\mathbf{p}=\mathbf{X}\boldsymbol{\beta},$$

where $\mathbf{p}=\left[\ln\left(\frac{p_i}{1-p_i}\right)\right]_{b\times 1}$,

$\mathbf{X}=\left[x_{ij}\right]_{b\times(m+1)}$ and

$\boldsymbol{\beta}=\left[\beta_j\right]_{(m+1)\times 1}$.

2.3.2.2 Method of estimating parameters

The main aim of the logistic regression model is to estimate the $m+1$ unknown parameters of $\boldsymbol{\beta}$. The method usually involved in the parameter estimations is the maximum likelihood estimations which involves finding the parameters for which the likelihood of the observed data in highest. This method is derived from the probability distribution function of the dependent variable.

Each y_i represents the number of binomial counts in the i^{th} group (category). Hence the joint probability density function f of Y_i is

$$f(p_i \mid \boldsymbol{\beta}) = \prod_{i=1}^{b} \frac{n_i!}{y_i(n_i - y_i)!} p_i^{y_i} (1 - p_i)^{n_i - y_i}, \tag{2.4}$$

where $f(p_i \mid \boldsymbol{\beta})$ describes $P(Y_i = y_i)$. There are $\binom{n_i}{y_i}$ different ways of arranging y_i successes from n_i trials for each group. Equation (2.4) is the joint probability density function expressing the values of $\mathbf{y}$ as a function of known fixed values of $\boldsymbol{\beta}$. The likelihood function differs from the probability density function only in the conditioning of the left hand side (White, 2007). Thus the likelihood function L expresses the values of $\boldsymbol{\beta}$ in terms of known fixed values of $\mathbf{y}$. Hence we have

$$L(\boldsymbol{\beta} \mid p_i) := f(p_i \mid \boldsymbol{\beta}) = \prod_{i=1}^{b} \frac{n_i!}{y_i(n_i - y_i)!} p_i^{y_i} (1 - p_i)^{n_i - y_i}. \tag{2.5}$$

We are looking for maximum likelihood estimates for $\boldsymbol{\beta}$ in equation (2.5). When the first derivative of a function is equated to 0, we get the critical points (maxima or minima). For a function of one variable, if the second derivative at the critical point is negative, then the critical point is maximum (Co, 2013). This is extended to the likelihood function. In order to find the maximum likelihood estimates, the first and second derivatives must be derived with respect to every element in $\boldsymbol{\beta}$. The factorial terms are independent of p_i and hence are regarded as constants. Equation (2.5) can be written as

$$L(\boldsymbol{\beta} \mid p_i) = \prod_{i=1}^{b} \left(\frac{p_i}{1 - p_i} \right)^{y_i} (1 - p_i)^{n_i}. \tag{2.6}$$

From equation (2.3), we set

$$\frac{p_i}{1-p_i} = \exp\left(\sum_{j=o}^{m} x_{ij}\beta_j\right), \tag{2.7}$$

and solving for p_i we get

$$p_i = \frac{\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}, \tag{2.8}$$

where exp is the exponential function.

Substituting equations (2.7) and (2.8), equation (2.6) becomes

$$L(\boldsymbol{\beta} \mid p_i) = \prod_{i=1}^{b}\left(\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)\right)^{y_i}\left(1-\frac{\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}\right)^{n_i}. \tag{2.9}$$

Equation (2.9) is made up of two products. It is evaluated to give:

$$L(\boldsymbol{\beta} \mid p_i) = \prod_{i=1}^{b}\left(\exp\left(y_i\sum_{j=0}^{m} x_{ij}\beta_j\right)\left(1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)\right)\right)^{-n_i}. \tag{2.10}$$

By the monotonicity of the logarithm function (Qi and Chen, 2007), the functional value of the likelihood function $L(\boldsymbol{\beta} \mid p_i)$ at $\boldsymbol{\beta}$ will be the same as the functional value of the log-likelihood function $l(\boldsymbol{\beta} \mid p_i)$ at $\boldsymbol{\beta}$. This allows us to apply natural logarithm to equation (2.10). Then we obtain

$$l(\boldsymbol{\beta} \mid p_i) := \ln L(\boldsymbol{\beta} \mid p_i) = \sum_{i=1}^{b} \left(y_i \left(\sum_{j=0}^{m} x_{ij}\beta_j \right) - n_i \ln \left(1 + \exp \left(\sum_{j=0}^{m} x_{ij}\beta_j \right) \right) \right). \qquad (2.11)$$

Equation (2.11) is called the log likelihood function. We find $\boldsymbol{\beta}$ by maximizing $l(\boldsymbol{\beta} \mid p_i)$.

Definition 2.1:

A maximum likelihood estimator (MLE) $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is a solution to the maximization problem in equation (2.11) which can be written as

$$\hat{\boldsymbol{\beta}} = \arg\max_{\Omega} l(\boldsymbol{\beta} \mid p_i) \qquad (2.12)$$

Assumptions (Gourieroux and Monfort, 1981)

To be able to find MLE of the log likelihood function, the following assumptions are made on the function.

i. The function $l(\boldsymbol{\beta} \mid p_i)$ is continuous on $\boldsymbol{\beta}$.

ii. The parameter space, Ω is compact.

iii. $\tilde{\boldsymbol{\beta}} \in Int(\Omega)$, where Ω is the compact parameter space and $Int(\Omega)$ is the interior of Ω.

The first two assumptions ensure the existence of $\hat{\boldsymbol{\beta}}$.

We find MLEs by considering the first and second partial derivatives of $l(\boldsymbol{\beta} \mid p_i)$ with respect to $\boldsymbol{\beta}$. The first derivative of the $l(\boldsymbol{\beta} \mid p_i)$ in equation (2.11) is given by

$$\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} = \sum_{i=1}^{b} \left(y_i x_{ij} - \frac{n_i}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)} \frac{\partial}{\partial \beta_j}\left(1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)\right)\right)$$

$$= \sum_{i=1}^{b} \left(y_i x_{ij} - \frac{n_i}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)} \exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right) \frac{\partial}{\partial \beta_j}\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)\right),$$

$j = 0,1,...,m.$

It is important to note that for each j,

$$\frac{\partial}{\partial \beta_j}\left(\sum_{j=0}^{m} x_{ij}\beta_j\right) = x_{ij}. \tag{2.13}$$

Then we have,

$$\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} = \sum_{i=1}^{b} \left(y_i x_{ij} - \frac{n_i}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)} \exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right) x_{ij}\right).$$

Substituting $p_i = \dfrac{\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}$, then we obtain

$$\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} = \sum_{i=1}^{b} \left(y_i x_{ij} - n_i p_i x_{ij}\right)$$

$$= \sum_{i=1}^{b} \left(y_i - n_i p_i\right) x_{ij}. \tag{2.14}$$

However, the expected value of each y_i, written as $\hat{y}_i$ is given by $\hat{y}_i = \mathrm{E}(y_i) = n_i p_i$ Therefore, equation (2.14) becomes

$$\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} = \sum_{i=1}^{b} \left(y_i - \hat{y}_i \right) x_{ij}. \tag{2.15}$$

Now let, $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, ..., \hat{y}_b)$ be a column vector with the same dimension as $\mathbf{y} = (y_1, y_2, ..., y_b)$. Equation (2.15) can be written in matrix form as

$$\left[\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} \right]_{j=0,1,...,m} = \mathbf{X}^T (\mathbf{y} - \hat{\mathbf{y}}). \tag{2.16}$$

The vector $\left[\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} \right]_{j=0,1,...,m}$ is also referred to as the Fisher Score vector and it is represented by $S(\boldsymbol{\beta} \mid p_i)$. When equated to $\mathbf{0}$, $S(\boldsymbol{\beta} \mid p_i)$ becomes a system of nonlinear multivariate equations. The MLE $\hat{\boldsymbol{\beta}}$ satisfies the system of nonlinear equations

$$S(\hat{\boldsymbol{\beta}} \mid p_i) := \left[\frac{\partial l(\hat{\boldsymbol{\beta}} \mid p_i)}{\partial \beta_j} \right]_{j=0,1,...,m} = \mathbf{0} \tag{2.17}$$

where $\mathbf{0}$ is a null vector. A solution to equation (2.17) will give the estimates of $l(\boldsymbol{\beta} \mid p_i)$ To make sure that these estimates are maximum values, we verify that the matrix of second partial derivatives (Hessian matrix) of $l(\boldsymbol{\beta} \mid p_i)$ is negative definite (Golub, 1996).

The matrix is formed by finding the second partial derivative of each of the equations in equation (2.14) with respect to each element denoted by β_k. The derivative of each element in $S(\boldsymbol{\beta} \mid p_i)$ is given by

$$\frac{\partial^2 l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k \partial \beta_j} = \frac{\partial}{\partial \beta_k}\left(\sum_{i=1}^{b} y_i x_{ij} - n_i x_{ij} p_i\right)$$

$$= \frac{\partial}{\partial \beta_k}\left(\sum_{i=1}^{b} y_i x_{ij}\right) - \frac{\partial}{\partial \beta_k}\left(\sum_{i=1}^{b} n_i x_{ij} p_i\right)$$

$$= -\frac{\partial}{\partial \beta_k}\left(\sum_{i=1}^{b} n_i x_{ij} p_i\right).$$

Since $p_i = \dfrac{\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}$, then we obtain

$$\frac{\partial^2 l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k \partial \beta_j} = -\sum_{i=1}^{b}\left(n_i x_{ij} \frac{\partial}{\partial \beta_k}\left(\frac{\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_{ij}\beta_j\right)}\right)\right), \tag{2.18}$$

where $j = 0,1,...,m$ and $k = 0,1,...,m$.

Using exponential differentiation and quotient rule readily gives the solution to equation (2.14) presented as:

$$l''(\boldsymbol{\beta}) := \frac{\partial^2 l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k \beta_j} = -\sum_{i=1}^{b}\left(n_i x_{ij} p_i (1-p_i) x_{ik}\right), \tag{2.19}$$

where $j = 0,1,...,m$ and $k = 0,1,...,m$.

Now let $\mathbf{W}$ be a $b \times b$ diagonal matrix with elements $w_i = n_i p_i (1-p_i)$. Thus $\mathbf{W}$ has the form

$$\mathbf{W} = \begin{bmatrix} n_1 p_1(1-p_1) & 0 & \cdots & 0 \\ 0 & n_2 p_2(1-p_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_b p_b(1-p_b) \end{bmatrix}$$

From equation (2.19) the Hessian matrix $\mathbf{H}(\boldsymbol{\beta} \mid p_i)$ can be written as

$$\mathbf{H}(\boldsymbol{\beta} \mid p_i) := \left[\frac{\partial^2 l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k \beta_j} \right] = -\mathbf{X}^T \mathbf{W} \mathbf{X} \tag{2.20}$$

Theorem 2.2 (Co, 2013)

Let $l(\boldsymbol{\beta} \mid p_i)$ be a multivariable function that is twice differentiable on $\boldsymbol{\beta}$. The vector $\hat{\boldsymbol{\beta}}$ yields local maximum values for $l(\boldsymbol{\beta} \mid p_i)$ if

$$S(\boldsymbol{\beta} \mid p_i) := \left. \frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = \mathbf{0} \text{ and } \mathbf{U}^T \mathbf{H}(\boldsymbol{\beta} \mid p_i)\mathbf{U} < 0,$$

for an arbitrary column matrix $\mathbf{U} \neq \mathbf{0}$, where $\mathbf{0}$ is a null vector.

Definition 2.3: (Co, 2013)

The Hessian matrix $\mathbf{H}(\boldsymbol{\beta} \mid p_i)$ is said to be negative definite if for any arbitrary positive column matrix $\mathbf{U}$,

$$\mathbf{U}^T \mathbf{H}(\boldsymbol{\beta} \mid p_i)\mathbf{U} < \mathbf{0}.$$

Next we make sure that the solution to equation (2.14) yields maximum values for $l(\boldsymbol{\beta} \mid p_i)$ by showing that $\mathbf{H}(\boldsymbol{\beta} \mid p_i)$ is negative definite. By applying Theorem 2.2, we can show that $\mathbf{H}(\boldsymbol{\beta} \mid p_i)$ is negative definite.

To apply Theorem 2.2, it is important to show that the $\mathbf{H}(\boldsymbol{\beta} \mid p_i)$ in equation (2.20) is negative semi-definite.

For an arbitrary column matrix $\mathbf{U} = \left[u_j \right]_{(m+1)\times 1}$,

$$\mathbf{U}^T \mathbf{H}(\boldsymbol{\beta} \mid p_i)\mathbf{U} = -\mathbf{U}^T \mathbf{X}^T \mathbf{W}\mathbf{X}\mathbf{U} = -\sum_{i=1}^{b}\sum_{j=0}^{m}(x_{ij}u_j)^2 w_i$$

w_i is always positive, therefore, $\mathbf{U}^T \mathbf{H}(\boldsymbol{\beta} \mid p_i)\mathbf{U} < \mathbf{0}$. Hence $\mathbf{H}(\boldsymbol{\beta} \mid p_i)$ is negative definite. Therefore, we have verified that the parameters $\boldsymbol{\beta}$ of $l(\boldsymbol{\beta} \mid p_i)$ are maximum likelihood parameters.

2.3.2.3 The Fisher Scoring iterative method

Setting each of the Score functions to 0,

$$\sum_{i=1}^{b}\left(y_i x_{ij} - n_i p_i x_{ij} \right) = 0,$$

gives a system of $m+1$ nonlinear equations with $m+1$ unknowns.

The solution is a vector with β_j elements. We have verified that the matrix of second partial derivatives is negative definite which means that the functional value of $l(\boldsymbol{\beta} \mid p_i)$ at $\boldsymbol{\beta}$ is maximum, thus, we can conclude that this vector contains parameter estimates for which each observed data has the highest chances of occurrence.

However, solving a system of nonlinear equations is not easy. This system can be considered as an unconstrained optimization problem. Most literature suggest the use of the Newton Raphson's algorithm to solving such systems of nonlinear equations (More and Cosnard, 1979; Givens and Hoeting, 2005).

Let $\boldsymbol{\beta}^{(0)}$ be the vector of initial approximations for each β_j, then the Newton Raphson's method for multivariate systems is given by

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \left[\mathbf{H}\left(\boldsymbol{\beta}^{(t)}\right)\right]^{-1} l'\left(\boldsymbol{\beta}^{(t)}\right), \tag{2.21}$$

where $t = 0,1,...,t_{\max}$ is the number of iterations and $t_{\max}$ is the maximum number of iterations

However, when considering likelihood functions, the Fisher Scoring algorithm has specially been designed to maximize such functions (van den Bos, 2007). The Fisher scoring algorithm is similar to the Newton's method. In the Fisher scoring algorithm, the additive inverse of the Hessian matrix is approximated by a Fisher information matrix. The Fisher information matrix is represented by

$$\mathbf{I}(\boldsymbol{\beta} \mid p_i) = -\mathrm{E}\left[\mathbf{H}(\boldsymbol{\beta} \mid p_i)\right]. \tag{2.22}$$

Using the Fisher information matrix to approximate the Hessian matrix, the Fisher scoring iterative method is given by

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \mathbf{I}^{-1}\left(\boldsymbol{\beta}^{(t)}\right) l'\left(\boldsymbol{\beta}^{(t)}\right), \tag{2.23}$$

where $t = 0,1,...,t_{\max}$.

From equations (2.16) and (2.20), the Fisher scoring method is given in matrix form by

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} + \mathrm{E}\left[\mathbf{X}^T\mathbf{W}\mathbf{X}\right]^{-1}\mathbf{X}^T(\mathbf{y}-\hat{\mathbf{y}}), \tag{2.24}$$

where $t = 0,1,...,t_{\max}$.

The Fisher scoring algorithm is preferred over the Newton Raphson algorithm because of two main advantages.

Firstly, under some regularity conditions (stated in section 2.4)

$$E\left[\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_j \partial\beta_k}\right] = -E\left[\left(\frac{\partial \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_j}\right)\left(\frac{\partial \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_k}\right)\right].$$

Hence the Fisher scoring algorithm requires only the first derivatives of the log-likelihood function in equation (11).

Secondly, the probability density function $f(p_i \mid \boldsymbol{\beta})$ is regular and the information matrix $\mathbf{I}(\boldsymbol{\beta} \mid p_i) = -\mathrm{E}\left[\mathbf{H}(\boldsymbol{\beta} \mid p_i)\right]$ is always positive definite, therefore, the Fisher information matrix is always invertible and prevents any chances of non-convergence unlike the Newton Raphson method (Garthwaite et al., 2002; Nielsen, 2003). This iterative procedure in equation (2.24) is linearly convergent which means that

$$\lim_{t\to\infty}\frac{\left\|\boldsymbol{\beta}^{(t+1)} - \hat{\boldsymbol{\beta}}\right\|}{\left\|\boldsymbol{\beta}^{(t)} - \hat{\boldsymbol{\beta}}\right\|} \le \eta,$$

for some $0 < \eta < 1$. One extra advantage of the fisher scoring method is that there is an inbuilt package in the R program that computers MLEs and their standard errors.

2.3.2.4 Sum Contrasts

In order to obtain the parameters of the categories that were assumed as reference groups, the reference groups were changed and the iterations will be repeated. The process involves creating a contrast matrix to change the reference group. Tongkumchum and McNeil (2009) have presented a method to create contrast matrices. This method also allows a comparison between dummy variables and the overall mean percentage.

2.3.2.5Contrast matrices

Tongkumchum and McNeil (2009), have proposed a method for constructing contrast matrices. However, for this thesis, a new version is used. This method is just a modification of what they presented and gives the same results. For a variable K with A categories used as explanatory variables, let the sample size in category a be n_a and the corresponding mean be $\bar{y}_a$ The overall mean percentage (mortality rate) is given by

$$\bar{y} = \frac{Total\ deaths}{Total\ observations} = \sum_{a=1}^{A} r_a \bar{y}_a \qquad (2.25)$$

where $r_a = \frac{n_a}{N}$ is the proportion of observations in category a. The contrast matrix C has a structure given as

$$C = \begin{bmatrix} I \\ r^* \end{bmatrix} \qquad (2.26)$$

I is an identity matrix with dimension $A-1$ and $r*$ is a row matrix with length $A-1$ and entries $\frac{-r_1}{r_A}, \frac{-r_2}{r_A}, ..., \frac{-r_{A-1}}{r_A}$. One column is deleted to obtain the desired contrast matrix which will correspond to the $A-1$ dummy variables specified in the regression model.

2.3.3 Adjusted probabilities

The value of each β_j obtained from equation (2.24) was then substituted into equation (2.3) to obtain the logistic regression predictive model. The value of β_j for $j = 0$ is known as the intercept and the value of β_j for $j > 0$ is known as the slope coefficient of the variable x_j.

The slope coefficients are interpreted as the effect of a unit of change in the variable x_j on the predictive model when all the other variables in the model are held constant. Based on the value of β_j a linear relationship between neonatal mortality and its risk factors will be established.

Once the parameters of the vector β were estimated, each β_j was substituted into equation (2.8) to estimate a new p_i, represented as $\hat{p}_i$. This $\hat{p}_i$ is known as fitted probability. The fitted probabilities $\hat{p}_i$ usually differ from the observed probabilities p_i and therefore, they were adjusted using a constant λ to make them approximately equal to the observed probabilities. The adjusted probabilities were then converted into proportions (percentages). This helped in making meaningful conclusions.

The Marquardt damping factor has been described as the most efficient way for adjusting proportions (Lourakis *et al*., 2005). The Marquardt damping factor was used to find adjusted proportions that corresponds to each p_i. The factor was used together with the Newton Raphson's iterative procedure. The adjusted proportions for a logistic regression has been presented by McNeil (2016), and it is given by

$$\hat{p}_i = \begin{cases} \dfrac{100}{1-\exp(-k-\beta_j)} & \beta_j \le 0, \\ \dfrac{100}{1+\exp(-k-\beta_j\lambda)} & \beta_j > 0, \end{cases} \tag{2.27}$$

$k = -\log\left(100/\bar{y} - 1\right)$. It remains to estimate λ. Each $\hat{p}_i$ must satisfy.

$$\sum_{i=1}^{b} n_i p_i = \sum_{i=1}^{b} n_i \hat{p}_i. \tag{2.28}$$

The constant λ was estimated using the Newton Raphson method with damping constant. The method of estimating λ is presented by McNeil (2016) as

$$\lambda_{t+1} = \lambda_t - \delta \frac{l(\boldsymbol{\beta} \mid p_i)}{l'(\boldsymbol{\beta} \mid p_i)}, \tag{2.29}$$

$\delta \in (0,1]$ is the damping factor and $t = 0,1,...t_{\max}$ is the number of iterations. The Newton's iterative method with damping factor in equation (2.29) computes the constant λ which

was used to adjust the fitted probabilities. An expected number of observations in each group was also be estimated. The model will be evaluated by plotting these adjusted probabilities against the observed probabilities.

2.4 Asymptotic Properties of the MLEs

To be sure that the MLEs are optimum solution to the likelihood function, it is important that their asymptotic properties be verified. The principles of maximum likelihood were first established in 1920 by R.A Fisher. He also established properties for maximizing the likelihood function $l(\boldsymbol{\beta} \mid p_i)$ (Myung, 2003; Aldrich, 1997).

Proposition 2.3. (Greene 2005).

Let the random variables Y_i be independent and identically distributed with density function $f(p_i \mid \boldsymbol{\beta})$ that satisfies all the regularity conditions. Then there exists a solution $\hat{\boldsymbol{\beta}}$ to the systems of nonlinear equations $S(\boldsymbol{\beta} \mid p_i) = \mathbf{0}$ and the solution has the following properties:

(a) $\hat{\boldsymbol{\beta}}$ is consistent for estimating $\boldsymbol{\beta}$;

(b) $\sqrt{n_i}(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})$ is asymptotically normal;

(c) $\hat{\boldsymbol{\beta}}$ is asymptotically efficient.

Next, we proceed to find the maximum likelihood estimates of equation (2.11). Before proceeding, we state and prove the regularity theorem of likelihood functions.

Theorem 2.4 (Albert and Anderson, 1984).

Let $f(p_i \mid \boldsymbol{\beta})$ be the probability density function of the binomial probability distribution. The MLEs of $f(p_i \mid \boldsymbol{\beta})$ exists and satisfies the asymptotic properties in proposition 2.3 if and only if $f(p_i \mid \boldsymbol{\beta})$ is regular. The conditions for regularity are:

C1 The probability distributions of the observations are distinct. Thus, if $\boldsymbol{\beta}1 \neq \boldsymbol{\beta}2,$ then $f(p_i \mid \boldsymbol{\beta}1) \neq f(p_i \mid \boldsymbol{\beta}2),$ for any two parameter vectors $\boldsymbol{\beta}1$ and $\boldsymbol{\beta}2$.

C2 The distributions have common support.

C3 Each random variable Y_i with probability density function is independent and identically distributed.

C4 The first and second derivatives of $\ln f(p_i \mid \boldsymbol{\beta})$ satisfy

$$E\left[\frac{\partial \ln\left(f(p_i \mid \boldsymbol{\beta})\right)}{\partial \beta_j}\right] = 0 \text{ and } \mathbf{I}(\boldsymbol{\beta} \mid p_i) = -E\left[\frac{\partial^2 \ln\left(f(p_i \mid \boldsymbol{\beta})\right)}{\partial \beta_k \partial \beta_j}\right].$$

C5 Since the Fisher information matrix $\mathbf{I}(\boldsymbol{\beta} \mid p_i)$ is a $(m+1)\times(m+1)$ covariance matrix and it is positive definite. We assume that each element in $\mathbf{I}(\boldsymbol{\beta} \mid p_i)$ is finite and that the matrix is positive definite for all $\boldsymbol{\beta} \in \Omega$ and $m+1$ statistics then

$$\frac{\partial}{\partial \beta_0} \ln f(p_i \mid \boldsymbol{\beta}), \frac{\partial}{\partial \beta_1} \ln f(p_i \mid \boldsymbol{\beta}), ..., \frac{\partial}{\partial \beta_j} \ln f(p_i \mid \boldsymbol{\beta}), \text{ for } j = 0,1,...m$$

are independent with probability 1

C6 The probability density function is three times differentiable and the third derivative is bounded by an integrable function $M(X)$. Thus

$$\left|\frac{\partial^3 \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_l \partial\beta_k \partial\beta_j}\right| \le M(X),$$

where $E\left(M(X)\right) < \infty$.

Proposition 2.5 Properties of the Fisher Score Vector (Edwards 1984)

Let $f(p_i \mid \boldsymbol{\beta})$ be a regular probability density function. Then,

1. $\mathbf{E}\left[S(\boldsymbol{\beta} \mid p_i)\right] = \int S(\boldsymbol{\beta} \mid p_i) f(p_i \mid \boldsymbol{\beta}) dp_i = \mathbf{0}$, where $\mathbf{0}$ is a null vector
2. $Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{E}\left[S(\boldsymbol{\beta} \mid p_i) S'(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{I}(\boldsymbol{\beta} \mid p_i)$, where $\mathrm{S}'(\boldsymbol{\beta} \mid p_i)$ is the transpose of $\mathrm{S}(\boldsymbol{\beta} \mid p_i)$.

Theorem 2.6: (Interchange of Integration and Differentiation)

Let $\ln f(p_i \mid \boldsymbol{\beta})$ be a differentiable function with respect to $\boldsymbol{\beta}$. If there exists a function $M(X) < \infty$ such that $\left|\frac{\partial \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_j}\right| \le M(X)$, then

$$\frac{\partial}{\partial\beta_j} \int \ln f(p_i \mid \boldsymbol{\beta}) dp_i = \int \frac{\partial}{\partial\beta_j} \ln f(p_i \mid \boldsymbol{\beta}) dp_i.$$

Chapter 3

Results

In this chapter, we prove that the likelihood function for the logistic regression is regular, and can be used to obtained parameter estimates $\hat{\beta}$ in theorem 2.4. The properties of the Fisher Score Vector are also discussed and the parameter estimates from the logistic regression model are presented for three different data.

3.1 Regularity of the likelihood function of the logistic regression model

C1: The probability distributions of the observations are distinct. Thus, if $\beta 1 \neq \beta 2$, then $p_{\beta 1} \neq p_{\beta 2}$, for any two parameter vectors $\beta 1$ and $\beta 2$. Let $\beta 1 = (\beta 1_0, \beta 1_1, ..., \beta 1_m)$ and $\beta 2 = (\beta 2_0, \beta 2_1, ..., \beta 2_m)$. We are going to show that if $\beta 1 \neq \beta 2$, then $p_{\beta 1} \neq p_{\beta 2}$.

We define the logistic regression model for all dummy variables x_j as

$$\ln\left(\frac{p_\beta}{1-p_\beta}\right) = \beta_0 + \beta_1 x_1 + ... + \beta_j x_j, \tag{3.1}$$

for $j = 0,1,2,...,m$. And we define

$$p_\beta = \frac{\exp\left(\sum_{j=o}^{m} x_j \beta_j\right)}{1+\exp\left(\sum_{j=o}^{m} x_j \beta_j\right)} \tag{3.2}$$

Let $\beta 1 = (\beta 1_0, \beta 1_1, ..., \beta 1_m)$ and $\beta 2 = (\beta 2_0, \beta 2_1, ..., \beta 2_m)$ be parameter vectors. Then for two different groups and for all dummy variables we have

$$p_{\beta 1} = \frac{\exp\left(\sum_{j=o}^{m} x_j \beta 1_j\right)}{1+\exp\left(\sum_{j=o}^{m} x_j \beta 1_j\right)}. \tag{3.3}$$

And

$$p_{\beta 2} = \frac{\exp\left(\sum_{j=o}^{m} x_j \beta 2_j\right)}{1+\exp\left(\sum_{j=o}^{m} x_j \beta 2_j\right)}. \tag{3.4}$$

Then if $\beta 1 = \beta 2$ then equations (3.3) and (3.4) are equal. This means that the probability density function is well defined. Next we show that if $\beta 1 \neq \beta 2$ then $p_{\beta 1} \neq p_{\beta 2}$.

Assume that $p_{\beta 1} = p_{\beta 2}$ then

$$\frac{\exp\left(\sum_{j=o}^{m} x_j \beta 1_j\right)}{1+\exp\left(\sum_{j=o}^{m} x_j \beta 1_j\right)} = \frac{\exp\left(\sum_{j=o}^{m} x_j \beta 2_j\right)}{1+\exp\left(\sum_{j=o}^{m} x_j \beta 2_j\right)}.$$

$$\exp\left(\sum_{j=0}^{m} x_j \beta 1_j\right)\left[1+\exp\left(\sum_{j=0}^{m} x_j \beta 2_j\right)\right] = \exp\left(\sum_{j=0}^{m} x_j \beta 2_j\right)\left[1+\exp\left(\sum_{j=0}^{m} x_j \beta 1_j\right)\right]. \tag{3.5}$$

The left hand side of equation (3.5) is given by

$$\exp\left(\sum_{j=0}^{m} x_j \beta 1_j\right)+\left[\exp\left(\sum_{j=0}^{m} x_j \beta 1_j\right)\right]\left[\exp\left(\sum_{j=0}^{m} x_j \beta 2_j\right)\right].$$

And the right hand side is given by

$$\exp\left(\sum_{j=0}^{m} x_j \beta 2_j\right) + \left[\exp\left(\sum_{j=0}^{m} x_j \beta 1_j\right)\right]\left[\exp\left(\sum_{j=0}^{m} x_j \beta 2_j\right)\right].$$

Both sides have a common term, hence equating the left hand side to the right hand side then we have

$$\exp\left(\sum_{j=0}^{m} x_j \beta 1_j\right) = \exp\left(\sum_{j=0}^{m} x_j \beta 2_j\right).$$

Taking natural logarithm of both side,

$$\sum_{j=0}^{m} x_j \beta 1 = \sum_{j=0}^{m} x_j \beta 2,$$

and then,

$$\sum_{j=0}^{m} \left(\beta 1_j - \beta 2_j\right) x_j = 0.$$

Since $\beta 1 \neq \beta 2$, then

$$\sum_{j=0}^{m} \left(\beta 1_j - \beta 2_j\right) \neq 0.$$

Therefore,

$$\sum_{j=0}^{m} x_j = 0.$$

However, since x_j are independent, we get a contradiction. Thus $p_{\beta 1} \neq p_{\beta 2}$. Therefore, this shows that the distributions are unique, if $\beta 1 \neq \beta 2$, then the observations are distinct.

C2: The distributions have common support. The logistic regression model for each of the b groups, is given by

$$p_i = \frac{\exp\left(\sum_{j=o}^{m} x_j \beta_j\right)}{1+\exp\left(\sum_{j=o}^{m} x_j \beta_j\right)}, \; i = 1,2,...,b\,.$$

The variables are $x_1, x_2, ..., x_m$ and the vector of parameters $\boldsymbol{\beta}$ takes the values $-\infty < \beta_j < \infty$. Each group follows the same model. Hence, the distributions have common support.

C3 Each random variable Y_i with probability density function $f(p_i \mid \boldsymbol{\beta})$ is independent and identically distributed. This is straight forward from the observations. Each group has an independent probability with density function $f(p_i \mid \boldsymbol{\beta})$.

C4 The first and second derivatives of $f(p_i \mid \boldsymbol{\beta})$ satisfy

$$\mathrm{E}\left(\ln f(p_i \mid \boldsymbol{\beta})\right) = 0 \text{ and } \mathbf{I}(\boldsymbol{\beta} \mid p_i) = -\mathrm{E}\left[\frac{\partial^2 \ln f(p_i \mid \beta)}{\partial \beta_k \partial \beta_j}\right].$$

The proof for condition C4 is exactly as in proposition 2.5 in section 3.2.

C5 Since the Fisher information matrix $\mathbf{I}(\boldsymbol{\beta} \mid p_i)$ is a $(m+1)\times(m+1)$ covariance matrix and it is positive definite. We assume that each element in $\mathbf{I}(\boldsymbol{\beta} \mid p_i)$ is finite and that the matrix is positive definite for all $\boldsymbol{\beta} \in \Omega$ and $m+1$ statistics then

$$\frac{\partial}{\partial \beta_0} \ln f(p_i \mid \boldsymbol{\beta}), \frac{\partial}{\partial \beta_1} \ln f(p_i \mid \boldsymbol{\beta}), ..., \frac{\partial}{\partial \beta_j} \ln f(p_i \mid \boldsymbol{\beta}), \text{ for } j = 0,1,...m$$

are independent with probability 1

To prove this condition, we take the first derivative of the logarithm of the probability density function for each group. This is given as

$$\frac{\partial \ln f(p|\boldsymbol{\beta})}{\partial \beta_j} = yx_j - \frac{nx_j \exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}, \quad j = 0,1,...,m,$$

where y is the number of observations of interest in the group and n is the total sample size of the group.

We write the vectors in $\frac{\partial \ln f(p|\boldsymbol{\beta})}{\partial \beta_j}$ in a way so that they are linearly dependent.

Then we have:

$$yx_j - \frac{nx_j \exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)} = \sum_{j=0}^{m} yx_j\pi_j - \sum_{j=0}^{m} nx_j \frac{\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}\pi_j.$$

This implies that

$$yx_j - \sum_{j=0}^{m} yx_j\pi_j - \frac{nx_j \exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)} + \sum_{j=0}^{m} nx_j \frac{\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}\pi_j = 0$$

Factorising common terms gives

$$y\left(x_j - \sum_{j=0}^{m} x_j\pi_j\right) - n\frac{\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j\beta_j\right)}\left(x_j - \sum_{j=0}^{m} x_j\pi_j\right) = 0$$

And $$\left(x_j - \sum_{j=0}^{m} x_j \pi_j\right)\left(y - n\frac{\exp\left(\sum_{j=0}^{m} x_j \beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j \beta_j\right)}\right) = 0$$

y and n are natural numbers, and the exponential function is irrational, therefore, it means that $y - n\frac{\exp\left(\sum_{j=0}^{m} x_j \beta_j\right)}{1+\exp\left(\sum_{j=0}^{m} x_j \beta_j\right)} \neq 0$. Thus,

$$x_j = \sum_{j=0}^{m} x_j \pi_j$$

Since the joint distribution for all x_j is continuous, the probability P_{β} of each x_j is

$$P_{\beta}(x_j) = \begin{cases} 1 & x_j \neq \sum_{j=0}^{m} \pi_j x_j \\ 0 & otherwise \end{cases}$$

This implies that the statistics are independent.

C6 The probability density function is three times differentiable and the third derivative is bounded by an integrable function $M(X)$. Thus

$$\left|\frac{\partial^3 \ln f(p_i \mid \beta)}{\partial\beta_l \partial\beta_k \partial\beta_j}\right| \leq M(X),$$

where $\mathbf{E}\left(M(X)\right) < \infty$.

Notice that the probability density function $f(p_i \mid \beta)$ is equal to the likelihood function $L(\beta \mid p_i)$, therefore, $\ln f(p_i \mid \beta) = l(\beta \mid p_i)$.

We find the first, second and third derivatives of the log-likelihood function. Since the first derivative is the score vector which is written as

$$l'(\boldsymbol{\beta} \mid p_i) = S(\boldsymbol{\beta} \mid p_i) = \sum_{i=1}^{b} \left(y_i x_{ij} - n_i p_i x_{ij} \right).$$

$$S(\boldsymbol{\beta} \mid p_i) = \sum_{i=1}^{b} \left(y_i x_{ij} - n_i \frac{\exp\left(\sum_{j=0}^{m} x_{ij} \beta_j \right)}{1 + \exp\left(\sum_{j=0}^{m} x_{ij} \beta_j \right)} x_{ij} \right).$$

The second partial derivative from equation (19) is

$$l''(\boldsymbol{\beta} \mid p_i) = \frac{\partial^2 l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k \partial \beta_j} = -\sum_{i=1}^{b} \left(n_i x_{ij} p_i (1 - p_i) x_{ik} \right).$$

Therefore, the third derivative with respect to β_l is

$$l'''(\boldsymbol{\beta} \mid p_i) = \frac{\partial^3 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_l \partial \beta_k \partial \beta_j} = -\sum_{i=1}^{b} \left(n_i x_{ij} x_{ik} p_i (1 - p_i) x_{il} (1 - p_i (1 - p_i) x_{ik}) \right).$$

$$\left| \frac{\partial^3 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_l \partial \beta_k \partial \beta_j} \right| = \left| -\sum_{i=1}^{b} n_i x_{ij} x_{ik} p_i (1 - p_i) x_{il} \left(1 - p_i (1 - p_i) x_{ik} \right) \right|$$

$$\leq \sum_{i=1}^{b} \left| n_i x_{ij} x_{ik} p_i (1 - p_i) x_{il} \left(1 - p_i (1 - p_i) x_{ik} \right) \right|$$

$$\leq \sum_{i=1}^{b} \left| n_i x_{ij} x_{ik} p_i (1 - p_i) x_{il} \right| \left| (1 - p_i (1 - p_i) x_{ik}) \right|$$

Since $\left| p_i (1 - p_i) \left(1 - p_i (1 - p_i) x_{ik} \right) \right| < 1,$ we obtain

$$\left|\frac{\partial^3 \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_l \partial\beta_k \partial\beta_j}\right| \le \sum_{i=1}^{b} n_i \left|x_{ij} x_{ik} x_{il}\right|$$

$$\le \sum_{i=1}^{b} n_i \left|x_{ij}\right|\left|x_{ik}\right|\left|x_{il}\right|$$

$$\le M(X),$$

and $E\big(M(X)\big) < \infty$.

We have verified that the probability density function is regular and, therefore, the MLE $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ exists and satisfies proposition 2.3. Next we proceed to find the parameters of the model using equation (2.23)

3.2 Properties of the Fisher Score Vector

Recall that the Score vector

$$S(\boldsymbol{\beta} \mid p_i) = \frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial\beta_j} = \mathbf{0},$$

where $\mathbf{0}$ is a null vector. For each $j = 0,1,2,...,m$

$$S(\boldsymbol{\beta} \mid p_i) = \left[\sum_{i=1}^{b} \frac{\partial \ln f(p_i \mid \boldsymbol{\beta})}{\partial\beta_j}\right]_{j=0,1,2,...,m} = \left[\sum_{i=1}^{b} S(\boldsymbol{\beta} \mid p_i)\right]_{j=0,1,2,...,m}.$$

The Hessian matrix is given by

$$\mathbf{H}(\boldsymbol{\beta} \mid p_i) = \begin{bmatrix} \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 (p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_j} \\ \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_1^2} & \cdots & \frac{\partial^2 (p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_j} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_k \partial \beta_0} & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_k \partial \beta_1} & \cdots & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_k \partial \beta_j} \end{bmatrix}_{(m+1)\times(m+1)}.$$

The Fisher Information matrix $\mathbf{I}(\boldsymbol{\beta} \mid p_i)$ as defined in equation (2.22) is the negative of the expectation of the Hessian matrix. Therefore, $\mathbf{I}(\boldsymbol{\beta} \mid p_i)$ is given by

$$\mathbf{I}(\boldsymbol{\beta} \mid p_i) = -E\begin{bmatrix} \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_j} \\ \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_1^2} & \cdots & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_j} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_k \partial \beta_0} & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_k \partial \beta_1} & \cdots & \frac{\partial^2 l(p_i \mid \boldsymbol{\beta})}{\partial \beta_k \partial \beta_j} \end{bmatrix}.$$

Then for all b groups,

$$\mathbf{I}(\boldsymbol{\beta} \mid p_i) = -E\left(\sum_{i=1}^{b} n_i x_{ij} p_i (1 - p_i) x_{ik} \right).$$

Properties of the Fisher Score Vector

1. $E\left[S(\boldsymbol{\beta} \mid p_i)\right] = \int S(\boldsymbol{\beta} \mid p_i) f(p_i \mid \boldsymbol{\beta}) dp_i = \mathbf{0}$, where $\mathbf{0}$ is a null vector.

By definition, the probability density function for each group $f(p_i \mid \boldsymbol{\beta})$ integrates to 1. Therefore,

$$\int f(p_i \mid \boldsymbol{\beta}) dp_i = 1.$$

The Fisher Score function for each $f(p_i \mid \boldsymbol{\beta})$ is given by $S(\boldsymbol{\beta} \mid p_i)$,

$$E\left[S(\boldsymbol{\beta} \mid p_i)\right] = \int S(\boldsymbol{\beta} \mid p_i) f(p_i \mid \boldsymbol{\beta}) dp_i.$$

But $S(\boldsymbol{\beta} \mid p_i) = \frac{\partial}{\partial \beta_j} \ln f(p_i \mid \boldsymbol{\beta})$, therefore,

$$E\left[S(\boldsymbol{\beta} \mid p_i)\right] = \int \frac{\partial \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j} f(p_i \mid \boldsymbol{\beta}) dp_i$$

$$= \int \frac{1}{f(p_i \mid \boldsymbol{\beta})} \frac{\partial}{\partial \beta_j} \ln f(p_i \mid \boldsymbol{\beta}) f(p_i \mid \boldsymbol{\beta}) dp_i.$$

Applying the theorem 2.6 and by C6, we have

$$E\left[S(\boldsymbol{\beta} \mid p_i)\right] = \frac{\partial}{\partial \beta_j} \int f(p_i \mid \boldsymbol{\beta}) dp_i$$

$$= \frac{\partial}{\partial \beta_j} (1)$$

$$= 0.$$

2 Next we show that $Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{E}\left[S(\boldsymbol{\beta} \mid p_i)S'(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{I}(\boldsymbol{\beta} \mid p_i)$, where $S'(\boldsymbol{\beta} \mid p_i)$ is the transpose of $S(\boldsymbol{\beta} \mid p_i)$.

The Score vector $S(\boldsymbol{\beta} \mid p_i)$ is given by $S(\boldsymbol{\beta} \mid p_i) = \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}$. The probability density function $f(p_i \mid \boldsymbol{\beta})$ is the same as $L(\boldsymbol{\beta} \mid p_i)$, hence, $\ln f(p_i \mid \boldsymbol{\beta}) = l(\boldsymbol{\beta} \mid p_i)$.

Now since $Var\left[\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}\right] = E\left[\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k}\right)\right]$, then we can write

$$Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{E}\left[S(\boldsymbol{\beta} \mid p_i)S'(\boldsymbol{\beta} \mid p_i)\right],$$

$$Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{E}\left[\begin{pmatrix} \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0} \\ \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1} \\ \vdots \\ \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j} \end{pmatrix} \begin{pmatrix} \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0} & \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1} & \cdots & \dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k} \end{pmatrix}\right].$$

This results in a square matrix given as

$$Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = E\begin{bmatrix} \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0}\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0}\right) & \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1}\right) & \cdots & \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k}\right) \\ \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0}\right) & \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1}\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1}\right) & \cdots & \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_0}\right) & \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}\right)\left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_1}\right) & \cdots & \left(\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}\dfrac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_k}\right) \end{bmatrix}, \qquad (3.6)$$

for $j = 0,1,2,...,m$ and $k = 0,1,2,...,m$.

The expected value of the matrix is equal to the expected value of each element in the matrix. The expected value of each element is given by

$$E\left(\frac{\partial l(\boldsymbol{\beta}\mid p_i)}{\partial \beta_j}\frac{\partial l(\boldsymbol{\beta}\mid p_i)}{\partial \beta_k}\right)=\int\left(\frac{\partial \ln f(p_i\mid\boldsymbol{\beta})}{\partial \beta_j}\frac{\partial \ln f(p_i\mid\boldsymbol{\beta})}{\partial \beta_k}\right)f(p_i\mid\boldsymbol{\beta})dp_i,$$

$$=\int\left(\left(\frac{1}{f(p_i\mid\boldsymbol{\beta})}\frac{\partial f(p_i\mid\boldsymbol{\beta})}{\partial \beta_j}\right)\left(\frac{1}{f(p_i\mid\boldsymbol{\beta})}\frac{\partial f(p_i\mid\boldsymbol{\beta})}{\partial \beta_k}\right)\right)f(p_i\mid\boldsymbol{\beta})dp_i.$$

Factorizing common terms gives

$$E\left(\frac{\partial l(\boldsymbol{\beta}\mid p_i)}{\partial \beta_j}\frac{\partial l(\boldsymbol{\beta}\mid p_i)}{\partial \beta_k}\right)=\int\frac{1}{f(p_i\mid\boldsymbol{\beta})}\left(\frac{\partial f(p_i\mid\boldsymbol{\beta})}{\partial \beta_j}\right)\left(\frac{\partial f(p_i\mid\boldsymbol{\beta})}{\partial \beta_k}\right)dp_i. \qquad (3.7)$$

We let $F_{\boldsymbol{\beta}_k}=\frac{\partial f(p_i\mid\boldsymbol{\beta})}{\partial \beta_k}$ and $F_{\boldsymbol{\beta}_j}=\frac{\partial f(p_i\mid\boldsymbol{\beta})}{\partial \beta_j}$. Then we can write equation (3.6) as,

$$Var\left[S(\boldsymbol{\beta}\mid p_i)\right]=\begin{bmatrix}\int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_0}F_{\boldsymbol{\beta}_0}dp_i & \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_0}F_{\boldsymbol{\beta}_1}dp_i & \cdots & \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_0}F_{\boldsymbol{\beta}_k}dp_i\\ \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_1}F_{\boldsymbol{\beta}_0}dp_i & \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_1}F_{\boldsymbol{\beta}_1}dp_i & \cdots & \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_1}F_{\boldsymbol{\beta}_k}dp_i\\ \vdots & \vdots & \ddots & \vdots\\ \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_j}F_{\boldsymbol{\beta}_0}dp_i & \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_j}F_{\boldsymbol{\beta}_1}dp_i & \cdots & \int\frac{1}{f(p_i\mid\boldsymbol{\beta})}F_{\boldsymbol{\beta}_j}F_{\boldsymbol{\beta}_k}dp_i\end{bmatrix}_{(m+1)\times(m+1)}$$

Now, we show that the expectation of the Hessian matrix $\mathbf{H}(\boldsymbol{\beta}\mid p_i)$ is the same as the matrix above. Recall that

$$\mathbf{H}(\boldsymbol{\beta} \mid p_i) = \begin{bmatrix} \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_k} \\ \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_1^2} & \cdots & \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_0} & \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_1} & \cdots & \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_k} \end{bmatrix} \text{ and}$$

Expectation of a matrix is equal to the expectation of each element in the matrix. Hence,

$$E\left[\mathbf{H}(\boldsymbol{\beta} \mid p_i)\right] = \begin{bmatrix} E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_0^2}\right) & E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1}\right) & \cdots & E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_0 \partial \beta_k}\right) \\ E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0}\right) & E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_1^2}\right) & \cdots & E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_1 \partial \beta_k}\right) \\ \vdots & \vdots & \ddots & \vdots \\ E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_0}\right) & E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_1}\right) & \cdots & E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_k}\right) \end{bmatrix}.$$

Each element in $E\left[\mathbf{H}(\boldsymbol{\beta} \mid p_i)\right]$ is evaluated as

$$E\left(\frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_k}\right) = \int \frac{\partial^2 \ln f(p_i \mid \boldsymbol{\beta})}{\partial \beta_j \partial \beta_k} f(p_i \mid \boldsymbol{\beta}) dp_i,$$

$$= \int \frac{\partial}{\partial \beta_j}\left(\frac{1}{f(p_i \mid \boldsymbol{\beta})} \frac{\partial f(p_i \mid \boldsymbol{\beta})}{\partial \beta_k}\right) f(p_i \mid \boldsymbol{\beta}) dp_i.$$

Applying the product rule of differentiation to the right hand side gives

$$E\left(\frac{\partial^2 \ln f(p_i|\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}\right)=\int\left[\frac{\partial}{\partial\beta_j}\left(\frac{1}{f(p_i|\boldsymbol{\beta})}\right)\left(\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_k}\right)+\frac{1}{f(p_i|\boldsymbol{\beta})}\frac{\partial^2 f(p_i|\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}\right]f(p_i|\boldsymbol{\beta})dp_i.$$

$$=\int\left[\left(-\frac{1}{f(p_i|\boldsymbol{\beta})^2}\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_j}\right)\left(\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_k}\right)+\frac{1}{f(p_i|\boldsymbol{\beta})}\frac{\partial^2 f(p_i|\boldsymbol{\beta})}{\partial\beta_k\partial\beta_j}\right]f(p_i|\boldsymbol{\beta})dp_i.$$

$$=\int-\frac{1}{f(p_i|\boldsymbol{\beta})}\left(\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_j}\right)\left(\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_k}\right)dp_i+\int\frac{\partial^2 f(p_i|\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}dp_i.$$

Now by Theorem 2.6, we can interchange differential and integral. Hence we have

$$\int\frac{\partial^2 f(p_i|\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}dp_i=\frac{\partial^2}{\partial\beta_j\partial\beta_k}\int f(p_i|\boldsymbol{\beta})dp_i=0.$$

Therefore,

$$E\left[\frac{\partial^2 \ln f(p_i|\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}\right]=\int-\frac{1}{f(p_i|\boldsymbol{\beta})}\left(\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_j}\right)\left(\frac{\partial f(p_i|\boldsymbol{\beta})}{\partial\beta_k}\right)dp_i.$$

Thus,

$$E\left[\frac{\partial^2 \ln f(p_i|\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}\right]=\int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_j}F_{\beta_k}dp_i.$$

Hence we have

$$E\left[\mathbf{H}(\boldsymbol{\beta}|p_i)\right]=\begin{bmatrix}\int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_0}F_{\beta_0}dp_i & \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_0}F_{\beta_1}dp_i & \cdots & \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_0}F_{\beta_k}dp_i\\ \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_1}F_{\beta_0}dp_i & \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_1}F_{\beta_1}dp_i & \cdots & \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_1}F_{\beta_k}dp_i\\ \vdots & \vdots & \ddots & \vdots\\ \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_j}F_{\beta_0}dp_i & \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_j}F_{\beta_1}dp_i & \cdots & \int-\frac{1}{f(p_i|\boldsymbol{\beta})}F_{\beta_j}F_{\beta_k}dp_i\end{bmatrix}.$$

It follows that

$$Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = -E\left[\mathbf{H}(\boldsymbol{\beta} \mid p_i)\right].$$

Hence, we have verified that

$$Var\left[S(\boldsymbol{\beta} \mid p_i)\right] = -E\left[\mathbf{H}(\boldsymbol{\beta} \mid p_i)\right] = \mathbf{I}(\boldsymbol{\beta} \mid p_i).$$

3.3 Asymptotic normality of MLE

The probability density function used in this thesis is regular, hence, from proposition 2.3, the MLEs obtained from the logistic regression models are asymptotically normal. Asymptotic normality means that, under regularity conditions, the MLE for each random variable approaches normal distribution and the convergence rate is $\sqrt{n_i}$.

Definition 3.1

Let $f(p_i \mid \boldsymbol{\beta})$ be a probability density function and let $l(p_i \mid \boldsymbol{\beta}) = \ln f(p_i \mid \boldsymbol{\beta})$ be the corresponding log-likelihood function. There exists a true but unknown vector $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_0, \tilde{\beta}_1, ..., \tilde{\beta}_m)$ which is a maximizer of the likelihood function (Gourieroux and Monfort, 1981). This is written as

$$\tilde{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}} \mathrm{E}\left(l(\boldsymbol{\beta} \mid p_i)\right)$$

Theorem 3.2 (Pawitan, 2001).

Let $f(p_i \mid \boldsymbol{\beta})$ be a probability density function. If $f(p_i \mid \boldsymbol{\beta})$ is regular and if $\hat{\boldsymbol{\beta}}$ is the MLE to $l(\boldsymbol{\beta} \mid p_i)$, then $\sqrt{n_i}(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) \to N\left(0, \frac{1}{\mathrm{I}(\boldsymbol{\beta} \mid p_i)}\right)$.

Thus, the distribution assumes normality as $n_i \to \infty$.

Proof:

At the MLE $\hat{\boldsymbol{\beta}}$, the first derivative of the log-likelihood function gives a null vector. Thus,

$$l'(\hat{\boldsymbol{\beta}}) := \left.\frac{\partial l(\boldsymbol{\beta} \mid p_i)}{\partial \beta_j}\right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = \mathbf{0}. \tag{3.8}$$

We can write the Taylor series expansion around the true parameter vector $\tilde{\boldsymbol{\beta}}$. This gives

$$\mathbf{0} := l'(\hat{\boldsymbol{\beta}}) = l'(\tilde{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})l''(\tilde{\boldsymbol{\beta}}) + \frac{(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})^2}{2} l'''(\tilde{\boldsymbol{\beta}}) + \ldots \tag{3.9}$$

By the Mean Value Theorem (MVT), we can truncate the Taylor series expansion in equation (3.9) at the second term. Then we have

$$\mathbf{0} = l'(\tilde{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})l''(\tilde{\boldsymbol{\beta}}). \tag{3.10}$$

Rearranging and multiplying both sides of equation (3) by $\sqrt{n_i}$ we obtain

$$\sqrt{n_i}(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) = \frac{-\sqrt{n_i} l'(\tilde{\boldsymbol{\beta}})}{l''(\tilde{\boldsymbol{\beta}})}. \tag{3.11}$$

However, $l''(\bar{\boldsymbol{\beta}}) \to \mathbf{H}(\boldsymbol{\beta})$, by the law of large numbers (Pawitan, 2001). Hence equation (3.11) can be written as

$$\sqrt{n_i}(\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\beta}}) = -\frac{\sqrt{n_i}\,l'(\tilde{\boldsymbol{\beta}})}{\left[\mathbf{H}(\bar{\boldsymbol{\beta}})\right]}, \tag{3.12}$$

where $\bar{\boldsymbol{\beta}}$ is a point between $\tilde{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$.

The numerator on the right hand side of equation (3.11) is given as,

$$\sqrt{n_i}\,l'(\tilde{\boldsymbol{\beta}}) = \sqrt{n_i}\left(\frac{1}{n_i}\sum_{i=1}^{b} l'(\tilde{\boldsymbol{\beta}} \mid p_i) - \mathbf{0}\right), \tag{3.13}$$

where $\frac{1}{n_i}\sum_{i=1}^{n} l'(\tilde{\boldsymbol{\beta}} \mid p_i)$ is the first derivative of the normalized log-likelihood function (Green, 2005). Recall that the expectation of the Fisher Score vector is a Null vector $\mathbf{0}$. Hence we can write

$$\sqrt{n_i}\,l'(\tilde{\boldsymbol{\beta}}) = \sqrt{n_i}\left(\frac{1}{n_i}\sum_{i=1}^{b} l'(\tilde{\boldsymbol{\beta}} \mid p_i) - \mathrm{E}\left(l'(\tilde{\boldsymbol{\beta}} \mid p_i)\right)\right). \tag{3.14}$$

Applying the Central Limit Theorem (Green 2005), equation (3.13) converges to normal distribution. Thus

$$\sqrt{n_i}\left(\frac{1}{n_i}\sum_{i=1}^{b} l'(\tilde{\boldsymbol{\beta}} \mid p_i) - \mathrm{E}\left(l'(\tilde{\boldsymbol{\beta}} \mid p_i)\right)\right) \to N\left(\mathbf{0}, Var(l'(\boldsymbol{\beta} \mid p_i))\right). \tag{3.15}$$

Combining equation (3.12) and equation (3.15), then we can write equation (3.11) as

$$-\frac{\sqrt{n_i}\,l'(\tilde{\boldsymbol{\beta}})}{l''(\tilde{\boldsymbol{\beta}})} \to N\left(\mathbf{0}, \frac{Var(l'(\boldsymbol{\beta} \mid p_i)}{\mathbf{I}(\boldsymbol{\beta} \mid p_i)^2}\right).$$

From preposition 2.5, $Var(l'(\boldsymbol{\beta} \mid p_i)) = \mathbf{I}(\boldsymbol{\beta} \mid p_i)$. Hence we have

$$\sqrt{n_i}(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) \to N\left(0, \frac{1}{\mathrm{I}(\boldsymbol{\beta} \mid p_i)}\right).$$

Therefore, the distribution converges to normal distribution and $\sqrt{n_i}$ is the convergence rate.

3.4 Preliminary Analysis and Results

A total of 5,195 babies were admitted to the neonatal unit from January 2013 to May 2014. Out of the 5,195 babies admitted to the unit, 3,022 (58.2%) were delivered at the KATH while 2,173 (41.8%) were delivered and referred from outside the KATH. More than half of the babies were males (55%). VLBW constituted 16.3% of the study sample while 60.5% of the babies were of normal weight. Majority of the babies (77.7%) had 5 minutes Apgar score within 8-10. By gestation, more than half (66.1%) had their gestational age missing and 13.7% were preterm. Prematurity was responsible for 24.7% of neonatal admissions followed by infections which accounted for 23.6%. Distribution of neonatal characteristics according to place of delivery is presented in table 3.1.

Preliminary analysis of the study sample revealed the proportions of neonatal mortality in each variable. The analysis showed that 20.5% of females and 20.0% of males died during the period. The proportion of death among males that were referred was much higher (27.4%) than males that were delivered at KATH (14.7%). There were very high proportions of death among babies with 5 minute Apgar score of less than 4 (67.3%).

Referred babies with 5 minute Apgar less than 4 recorded 61.1% of mortality while that for non-referred babies was 70.7%. There was an overall 32.7% of death among babies with 5 minute Apgar score 4-7; babies referred had 42% of neonatal death and babies not referred had 28% of neonatal death.

Table 3.1 Distributions of the variables

Variables	Survived (%)	Died (%)	p-value
Gender			0.7186
Female	1860 (79.5)	480 (20.5)	
Male	2282 (80)	573 (20)	
Birth Weight			< 0.01
VLBW	431 (50.9)	416 (49.1)	
LBW	1029 (85.3)	178 (14.7)	
Normal	2682 (85.4)	459 (14.6)	
5 Minute Apgar score			< 0.01
0-3	66 (32.7)	136 (67.3)	
4-7	644 (67.3)	313 (32.7)	
8-10	3432 (85)	604 (15)	
Gestational age			< 0.01
Term	963 (91.7)	87 (8.3)	
Preterm	487 (68.5)	224 (31.5)	
Not Stated	2692 (78.4)	742 (21.6)	
Delivery Mode			< 0.01
SVD	2358 (76.2)	736 (23.8)	
C/S	1736 (84.8)	311 (15.2)	
Vacuum	48 (88.9)	6 (11.1)	
Place of Delivery			< 0.01
KATH	2573 (85.1)	449 (14.9)	
Referred	1569 (72.2)	604 (27.8)	
Admission age			< 0.01
1 day	3210 (78.3)	888 (21.7)	
2-7 days	728 (85.1)	127 (14.9)	
8-14 days	102 (86.4)	16 (13.6)	
15-28 days	102 (82.3)	22 (17.7)	
Diagnosis			< 0.01
Prematurity	888 (69.1)	397 (30.9)	
Respiratory Distress	208 (72.2)	80 (27.8)	
Infections	1120 (91.7)	102 (8.3)	
Congenital anomalies	377 (81.1)	88 (18.9)	
Neonatal Jaundice	799 (91.8)	71 (8.2)	
Birth asphyxia	638 (67.5)	307 (32.5)	
Others	112 ((93.3)	8 (6.7)	

The proportion of death in VLBW was 49.1%. By gestation, preterm birth had the highest proportion of neonatal death (31.5%) while 21.6% of babies with missing gestation also died. According to the discharge diagnosis, birth asphyxia was responsible for more neonatal deaths. 32.5% of asphyxiated babies couldn't survive followed by prematurity, with a proportion of 30.9%.

Every analytic procedure that uses regression models is preceded by a comparison of each variable with the outcome of interest. The analysis is accompanied by a significance level. This is also univariate analysis. The univariate analysis helps to select the factors that will be used in the logistic regression analysis. Pearson's chi-square test was used to test for the association of each variable with the outcome (neonatal mortality) at 95% significance level. Thus, for a variable to be significantly associated with neonatal mortality, the p-value must be less than 0.01.

From the univariate analysis only sex showed no significant association with neonatal mortality among both referred babies and those delivered at KATH. Thus, it was excluded from the multivariate analysis. Other variables such as birth weight, 5 minute Apgar score, gestational age, delivery mode, admission age and main diagnosis were all significantly associated with neonatal mortality. Table 3.2 shows the proportion of neonatal deaths in each category and the p-values among babies delivered at KATH

The results from univariate analysis on babies delivered at KATH suggest that 5 minute Apgar score, birth weight, gestational age, delivery mode, admission age and discharge diagnosis are significant predictors of death.

Table 3.2 Proportions of mortality in each category with p-values (KATH).

Variables	Survived (%)	Died (%)	p-value
Gender			0.7768
Female	1170 (84.9)	208 (15.1)	
Male	1403 (85.3)	241 (14.7)	
5 Minute Apgar			< 0.01
0-3	38 (29.2)	92 (70.8)	
4-7	466 (71.7)	184 (28.3)	
8-10	2069 (92.3)	173 (7.70)	
Birth Weight			< 0.01
VLBW	221 (49.7)	224 (50.3)	
LBW	682 (91.2)	66 (8.80)	
Normal	1670 (91.3)	159 (8.70)	
Gestational age			< 0.01
Term	853 (93.3)	61 (6.7)	
Preterm	348 (70.3)	147 (29.7)	
Not Stated	1372 (85.1)	241 (14.9)	
Delivery Mode			< 0.01
SVD	1180 (81.9)	261 (18.1)	
C/S	1353 (88)	185 (12)	
Vacuum	40 (93)	3 (7.0)	
Admission age			< 0.01
1 day	2345 (84.2)	441 (15.8)	
2-7 days	199 (97.1)	6 (2.90)	
8-14 days	14 (93.3)	1 (6.7)	
15-28 days	15 (93.7)	15 (6.30)	
Discharge Diagnosis			< 0.01
Prematurity	556 (71.4)	223 (28.6)	
Respiratory Distress	169 (84.5)	31 (15.5)	
Infections	697 (96.3)	27 (3.70)	
Congenital anomalies	209 (90.1)	23 (9.90)	
Neonatal Jaundice	449 (97.0)	14 (3.00)	
Birth asphyxia	408 (76.1)	128 (23.9)	
Others	3 (3.40)	85 (96.6)	

The univariate analysis conducted on referred babies revealed that both sex and delivery mode are not significantly associated with neonatal death. These variables have p-values more than 0.01. They were, thus, excluded from the multivariate analysis. Table 3.3 presents the univariate results and proportions of neonatal mortality among babies referred to the neonatal unit.

3.5 Results from the logistic regression model

The logistic regression model was fitted based on the results from the univariate analysis. Variables that failed to meet significance from the univariate analysis were excluded from the logistic regression model. Thus the model was defined as in equation (2.3). The value of β_j for $j = 0$ is known as the intercept and the value of β_j for $j > 0$ is known as the slope coefficient of variable x_j. The slope coefficients are interpreted as the effect of a unit of change in the variable x_j on the predicted logits when all the other variables in the model are held constant. β_j were used to calculate the $\hat{p}_i$. If $\beta_j > 0$ for any dummy variable, it is means the association between that variable and the outcome is positive; if $\beta_j = 0$ it implies that there is no association between the variable and the outcome. When $\beta_j < 0$ it implies that there is a negative association between the variable and outcome.

Table 3.3 Proportions of mortality in each category with p-values (Referred).

Variables	Survived (%)	Died (%)	p-value
Gender			0.692
Female	690 (71.7)	272 (28.3)	
Male	879 (72.6)	332 (27.4)	
5 Minute Apgar			<0.01
0-3	28 (37.5)	44 (62.5)	
4-7	178 (58.0)	129 (42.0)	
8-10	1363 (76.0)	431 (24.0)	
Birth weight			< 0.01
VLBW	210 (52.2)	192 (47.8)	
LBW	347 (75.6)	112 (24.4)	
Normal	1012 (77.1)	300 (22.9)	
Gestational age			< 0.01
Term	110 (80.9)	26 (19.1)	
Preterm	139 (64.4)	77 (35.4)	
Not Stated	1320 (72.5)	501 (27.5)	
Delivery mode			0.22
SVD	1178 (71.3)	475 (28.7)	
C/S	383 (75.2)	126 (24.8)	
Vacuum	8 (80.0)	3 (20.0)	
Admission age			< 0.01
1 day	865 (65.9)	447 (34.1)	
2-7 days	529 (81.4)	121 (18.6)	
8-14 days	88 (85.4)	15 (14.6)	
15-28 days	87 (80.6)	21 (19.4)	
Discharge Diagnosis			< 0.01
Prematurity	332 (65.6)	174 (34.4)	
Respiratory Distress	39 (44.3)	49 (56.7)	
Infections	423 (84.9)	75 (15.1)	
Congenital anomalies	168 (72.1)	65 (27.9)	
Neonatal Jaundice	350 (86.0)	57 (14.0)	
Birth asphyxia	230 (56.2)	179 (43.8)	
Others	27 (84.4)	5 (15.6)	

3.5.1 Results from KATH deliveries

From the univariate analysis that was conducted to the babies delivered at KATH, only gender was not significantly associated with neonatal mortality. Thus, all the remaining 6 main variables were included in the logistic regression model. From the categorization, there were 23 dummy variables were obtained. The variables were arranged in the same order as shown in table 3.2. Using equation (2.24) and setting $\beta^{(0)}$ as the initial guess, the vector β of parameters is presented in Table 3.4

Based on the results from the logistic regression, birth weight, 5 minute Apgar score, gestational age, and discharge diagnosis had p-value <0.01, while admission age and delivery mode had p-values > 0.01.

Table 3.4 Parameter estimates from the logistic regression model (KATH)

Variable	Sample size	Coefficient	Standard Error	p-value
Intercept		-2.4686	0.0887	< 0.01
5 Minute Apgar				**<0.01**
0-3	130	2.6367	0.2278	< 0.01
4-7	650	0.7041	0.1113	< 0.01
8-10	2242	-0.3570	0.0368	< 0.01
Birth Weight				**<0.01**
VLBW	445	1.9847	0.1755	<0.01
LBW	748	-0.2692	0.1301	<0.01
Normal	1829	-0.3728	0.0754	<0.01
Gestational age				**<0.01**
Term	914	-0.4464	0.1212	<0.005
Preterm	495	0.3339	0.1433	<0.005
Not Stated	1614	0.1505	0.0622	0.262

Table 3.4 (Continued)

Variable	Sample size	Coefficient	Standard Error	p-value
Intercept		-2.4686	0.0887	< 0.01
Delivery mode				**0.689**
SVD	1441	-0.0299	0.0689	0.595
C/S	1538	0.0406	0.0652	<0.005
Vacuum	43	-0.4490	0.6751	0.478
Admission age				**0.994**
1 day	2786	0.0039	0.0313	0.562
2-7 days	205	-0.0790	0.4125	0.874
8-14 days	15	0.1465	1.1092	0.906
15-28 days	16	0.1965	1.0523	0.855
Discharge Diagnosis				**<0.01**
Prematurity	779	0.0415	0.1526	0.713
Respiratory Distress	200	1.0286	0.2193	0.055
Infections	724	-0.4811	0.1743	0.649
Congenital anomalies	232	0.6690	0.2398	0.186
Neonatal Jaundice	463	-1.0822	0.2597	0.184
Birth asphyxia	536	0.8847	0.1501	0.082
Others	88	-0.2048	0.5943	<0.01

This indicates that the association of admission age and delivery mode with neonatal mortality is very weak. A reduced logistic regression model is fitted using only the variables that are significant from the first model. Thus, 5 minute Apgar score, birth weight, gestation and discharge diagnosis are used in the reduced model. The results from the reduced model are presented in the table 3.5 below.

Table 3.5 Parameter estimates from the reduced model (KATH)

Variable	Sample size	Coefficient	Standard Error	p-value
Intercept		-2.4656	0.0887	< 0.01
5 Minute Apgar				**<0.01**
0-3	130	2.6120	0.2247	< 0.01
4-7	650	0.6870	0.1070	< 0.01
8-10	2242	-0.3511	0.0351	< 0.01
Birth Weight				**<0.01**
VLBW	445	1.9947	0.1747	<0.01
LBW	748	-0.2623	0.1299	0.043
Normal	1829	-0.3781	0.0751	<0.01
Gestational age				**<0.01**
Term	914	-0.4411	0.1207	<0.01
Preterm	495	0.3345	0.1431	0.019
Not Stated	1614	0.1473	0.0618	0.017
Discharge Diagnosis				**<0.01**
Prematurity	779	0.0436	0.1525	0.775
Respiratory Distress	200	1.0353	0.2190	<0.01
Infections	724	-0.4810	0.1741	<0.01
Congenital anomalies	232	0.6656	0.2389	<0.01
Neonatal Jaundice	463	-1.0943	0.2574	<0.01
Birth asphyxia	536	0.8874	0.1489	<0.01
Others	88	-0.1847	0.5930	0.755

Figure 3.1a displays the Q-Q plot of the residuals. The plot indicates that the logistic regression model has provided an acceptable fit for the outcome variable (neonatal mortality) and predictor variables (5 minute Apgar score, birth weight, gestation and discharge diagnosis) among babies delivered at KATH. The residuals lie along a straight line, which means that parameters are normally distributed. Figure 3.1b shows the plot of the expected and observed values. The values almost match each other in a vertical pattern.

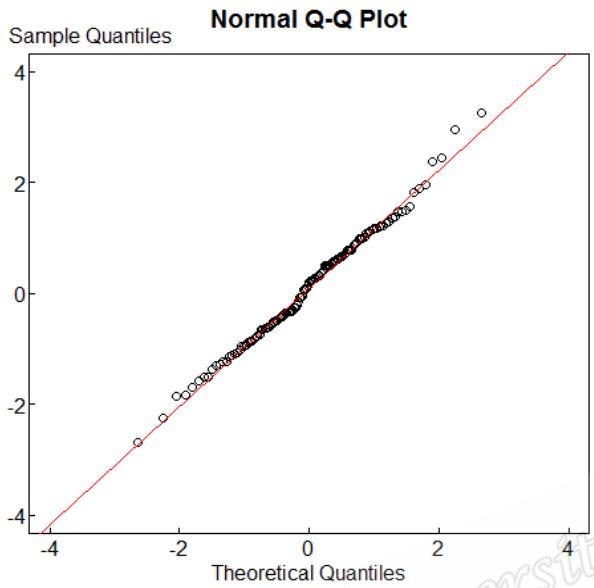

Figure 3.1a Normal quantile-quantile plot of residuals

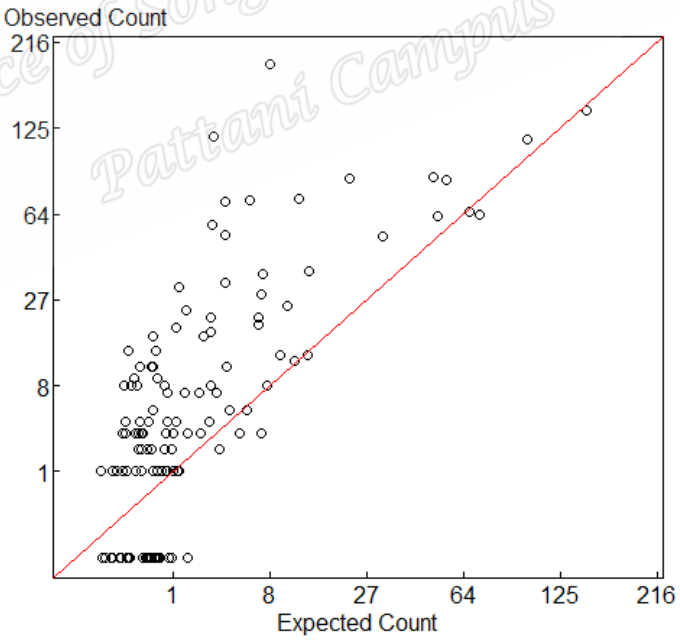

Figure 3.1b Expected and observed values

Fitted probabilities for each category were calculated from β. Observed and fitted probabilities of variables that were significant from the logistic regression model are presented in Fig 3.2 with the overall mortality rate of the babies delivered at KATH.

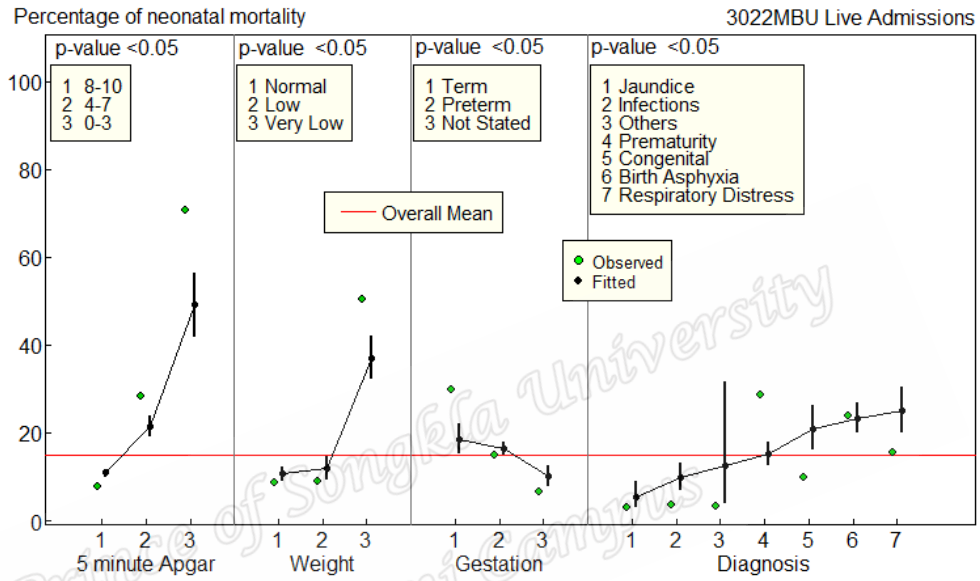

Figure 3.2 Confidence interval plot from the logistic regression model from the logistic regression model for babies delivered at KATH

The horizontal red line is the mean mortality rate.

A Receiver Operating Characteristic (ROC) curve (Figure 3.3) with corresponding ROC bar plot (Figure 3.4) of this model shows that the predictive power of the model (area under the curve) is 72.53%. This predictive power indicates that the logistic regression model for babies delivered at KATH and referred to the neonatal unit correctly predicts 72.53% of neonatal deaths. This prediction is based on all four variables, 5 minute Apgar score (apGrp), birth weight (wtGrp), discharge diagnosis (Diag) and gestational age (Ges).

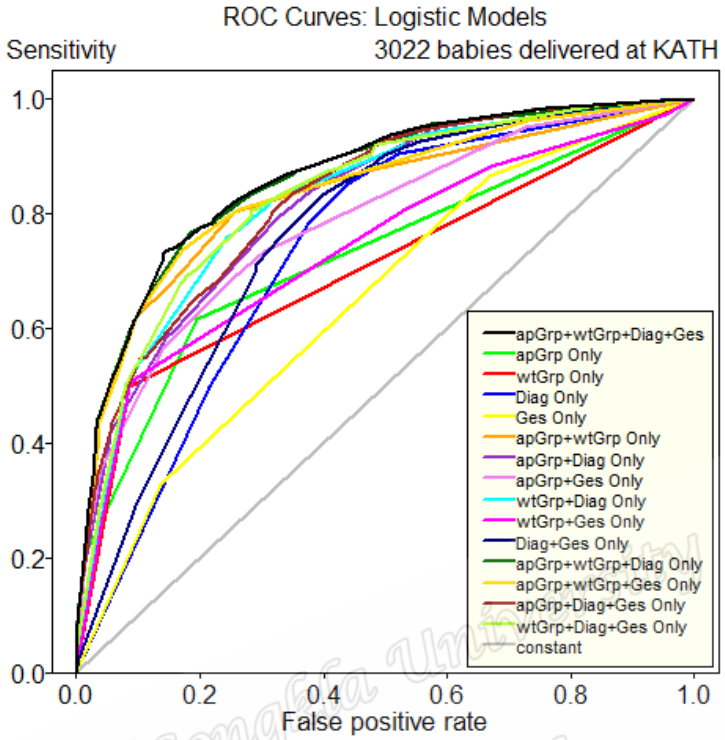

Figure 3.3 ROC curves for logistic regression model from KATH deliveries

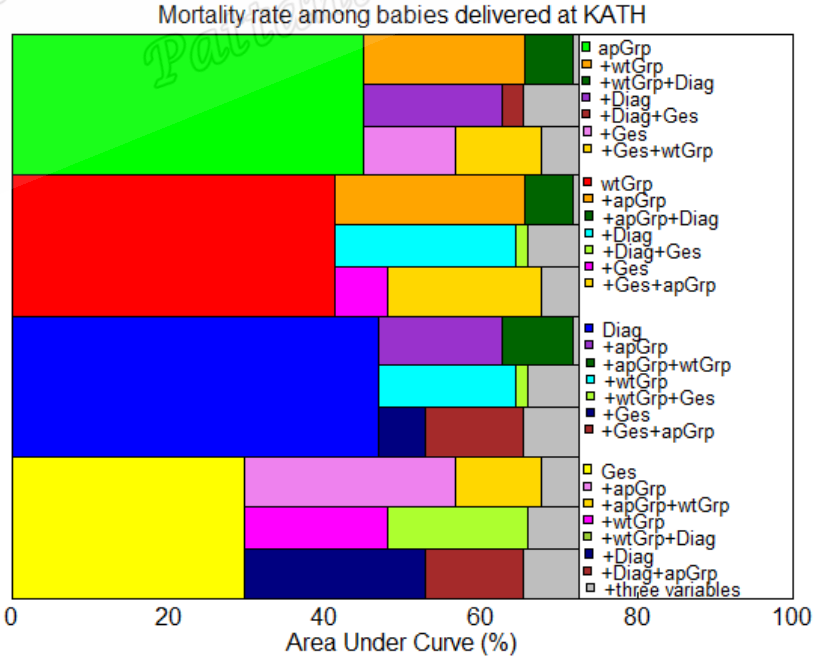

Figure 3.4 ROC bars for logistic regression model from KATH deliveries

The independent variables, 5 minute Apgar score (apGrp), birth weight (wtGrp), discharge diagnosis (Diag), and gestation (Ges) individually accounted for 45.00%, 41.30%, 46.96% and 29.76% of mortality respectively. When the variables are combined in pairs, 5 minute Apgar score and birth weight explains 65.69% of mortality and discharge diagnosis and gestation accounts for 66.05% mortality. A model with three variables, 5 minute Apgar score, birth weight and gestation explains 67.80% while a different model with (Figure 3.4). The ROC curve also shows that all the variables predict a significant level of neonatal mortality.

3.5.2 Results from Referred babies

After univariate analysis, the results revealed that gender and delivery mode were not significantly associated with neonatal death among babies referred to the neonatal unit. Therefore, birth weight, 5 minute Apgar score, gestational age, admission age and discharge diagnosis were included in the logistic regression model. The model was formulated with 20 dummy variables. The solution vector β of parameters is obtained as presented in table 3.6

Table 3.6 Parameter estimates from the logistic regression model (Referred)

Variable	Sample Size	Coefficient	Standard Error	p-value
Intercept		-1.8309	0.5261	<0.01
5 Minute Apgar				**0.01**
0-3	72	0.7757	0.2515	0.02
4-7	307	0.2444	0.1238	<0.01
8-10	1794	-0.0730	0.0239	< 0.01
Birth Weight				**<0.01**
VLBW	402	1.3281	0.1502	<0.01
LBW	459	-0.0623	0.1049	0.55
Normal	1312	-0.3851	0.0589	<0.01
Gestational age				**0.09**
Term	136	-0.4310	0.2252	0.04
Preterm	216	-0.0770	0.1595	0.62
Not Stated	1821	0.0413	0.0237	0.07
Admission age				**0.38**
1 day	1312	0.0531	0.0516	0.30
2-7 days	650	-0.0437	0.0955	0.492
8-14 days	103	-0.4489	0.2824	0.096
15-28 days	108	0.0453	0.2505	0.977
Discharge Diagnosis				**<0.01**
Prematurity	506	-0.4852	0.1406	0.662
Respiratory Distress	88	1.5113	0.2228	<0.01
Infections	498	-0.3648	0.1156	0.832
Congenital anomalies	233	0.4041	0.1481	0.195
Neonatal Jaundice	407	-0.4462	0.1371	0.713
Birth asphyxia	409	0.9532	0.1205	0.016
Others	32	-0.2577	0.4864	<0.01

After fitting the logistic regression model, gestational age and admission was found to be weakly associated with neonatal death. Their p-values were greater than 0.01. The significant variables were used to fit the final reduced model and subsequently used to create the plot. Parameter estimates from the reduced model are presented in table 3.7

Table 3.7 Parameter estimates from the reduced model (Referred)

Variable	Sample Size	Coefficient	Standard Error	p-value
Intercept		-1.0992	0.0545	<0.01
5 Minute Apgar				**<0.01**
0-3	72	0.7912	0.2506	<0.01
4-7	307	0.2463	0.1214	0.04
8-10	1794	-0.0740	0.0234	< 0.01
Birth Weight				**<0.01**
VLBW	402	1.3518	0.1469	<0.01
LBW	459	-0.0712	0.1040	0.49
Normal	1312	-0.3893	0.0573	<0.01
Discharge Diagnosis				**<0.01**
Prematurity	506	-0.4774	0.1387	<0.01
Respiratory Distress	88	1.4966	0.2196	<0.01
Infections	498	-0.3879	0.1130	<0.01
Congenital anomalies	233	0.3966	0.1471	<0.01
Neonatal Jaundice	407	-0.4565	0.1310	<0.01
Birth asphyxia	409	0.9865	0.1152	<0.01
Others	32	-0.2206	0.4854	0.65

From the reduced logistic regression analysis, all the variables remained significantly associated with neonatal mortality. Babies with VLBW, babies with Respiratory distress and babies with 5 minute Apgar score had very high risk of neonatal death and corresponding high mortality rates. Figure 3.5 shows the plot of observed and fitted probabilities of the significant variables.

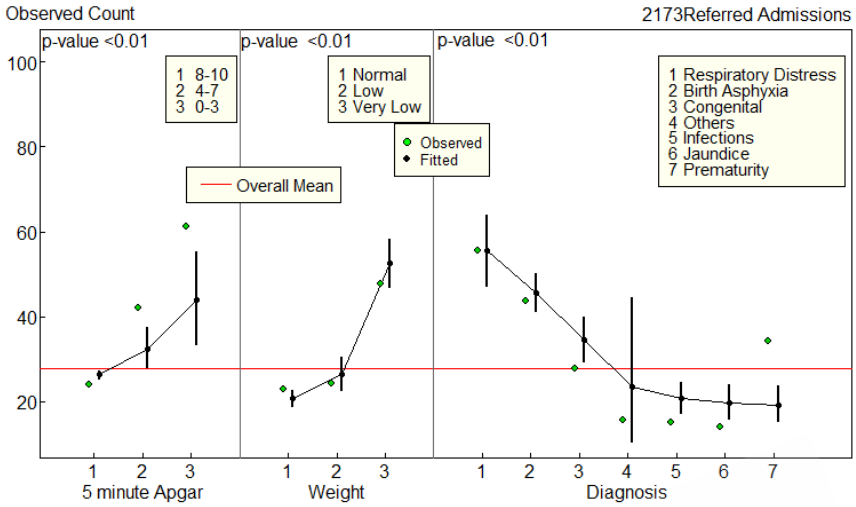

Figure 3.5 Observed and fitted probabilities from the reduced logistic regression model for babies referred to the neonatal unit

The Q-Q plot and the plot of expected vs. observed are presented in Figure 3.6a and 3.6b. All the residuals lie on the slope which means that they are normally distributed.

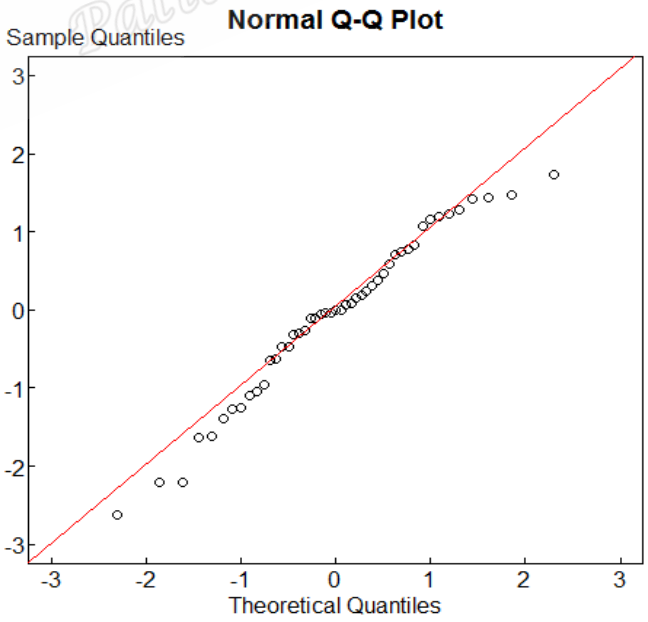

Figure 3.6a Normal quantile-quantile plot of residuals

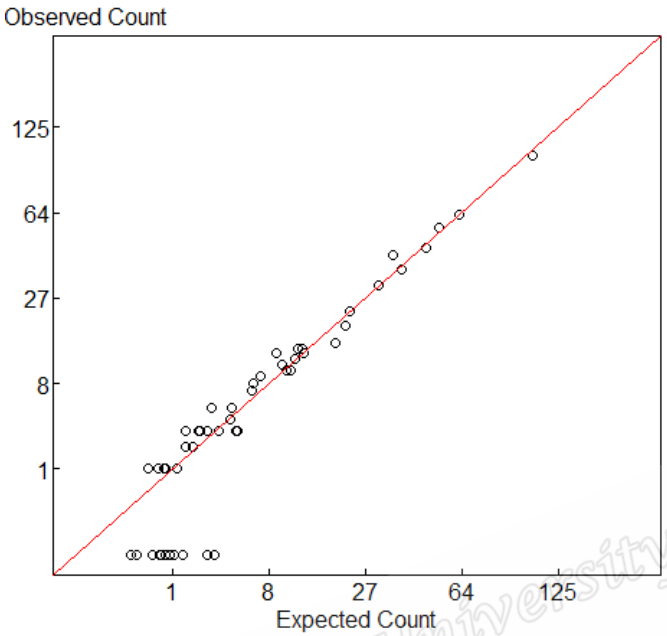

Figure 3.6b Expected and observed values

A similar ROC curve was generated for the model (Figure 3.7).

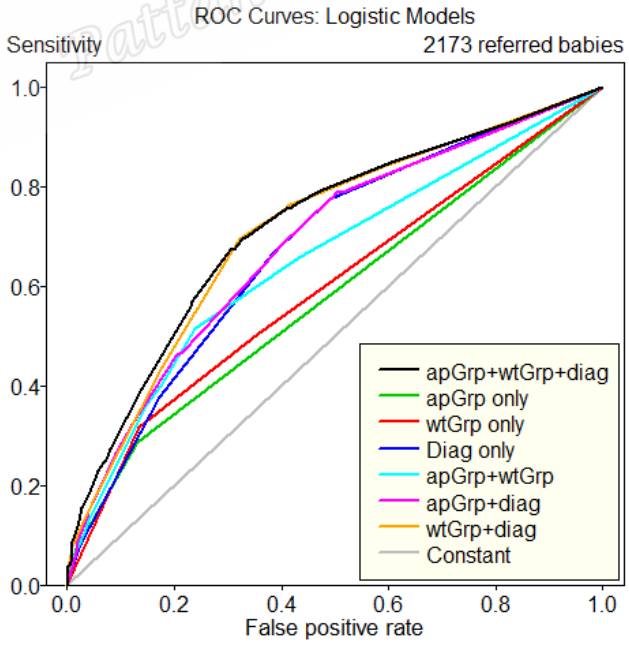

Figure 3.7 ROC curves for logistic regression model from referred deliveries

The three variables used 5 minute Apgar score, birth weight and discharge diagnosis could explain a very low variation in the mortality of babies referred to the unit. The area under the curve is 43.7. Each variable used in the model, Apgar score, birth weight and diagnosis explains 15.96%, 19.38% and 35.06% respectively. Combining two variables, birth weight and diagnosis, could account for 42.10% of mortality (Figure 3.8)

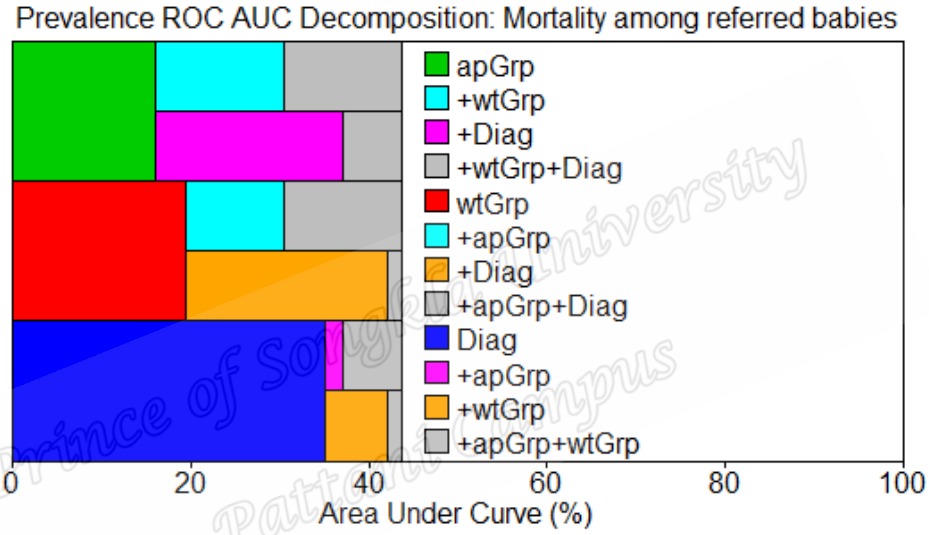

Figure 3.8 ROC bars for logistic regression model from referred babies.

3.5.3 Results from combined data

Univariate and multivariate analyses were repeated for the combined data. The univariate analysis showed that sex had no association with neonatal death. Place of delivery was included in logistic regression model; hence, the model was fitted with 7 variables which are categorized into 18 dummy variables. The solution vector of parameters obtained from the model is presented in Table 3.8. The multivariate analysis further revealed that admission age and delivery mode were not significantly associated

with the outcome. Babies having VLBW had 7 times higher risk of death and the percentage of mortality was also above the mean. Babies referred to the neonatal unit had a higher percentage of mortality. Babies with birth asphyxia, congenital anomalies, infections and respiratory distress had higher neonatal mortality than the overall mean.

To construct the reduced model, admission age and delivery mode were not included. The reduced model was fitted with 5 independent variables. The results revealed that all the five variables are significantly associated with neonatal death (Table 3.9)

The Q-Q plot presented in Figure 3.9 shows that the parameters are normally distributed. All the independent variables and dummy variables remained significantly associated with neonatal death. They were, thus, used to generate the confidence interval plot of proportions presented in Figure 3.10.

Table 3.8 Parameter estimates from the logistic regression model (Overall)

Variable	Sample Size	Coefficient	Standard Error	p-value
Intercept		-2.77054	0.39643	<0.01
5 Minute Apgar				**<0.01**
0-3	202	1.8982	0.1693	<0.01
4-7	957	0.5057	0.0798	<0.01
8-10	4036	-0.2149	0.0214	< 0.01
Birth Weight				**<0.01**
VLBW	847	1.6085	0.1123	<0.01
LBW	1207	-0.1646	0.0793	0.04
Normal	3141	-0.3705	0.0452	<0.01
Gestational age				**<0.01**
Term	1050	-0.5545	0.1072	<0.01
Preterm	711	0.2492	0.1030	0.02
Not Stated	3434	0.1180	0.0332	<0.01
Delivery mode				**0.44**
SVD	3094	0.0189	0.0356	0.76
C/S	2047	-0.0160	0.0537	0.60
Vacuum	54	-0.4769	0.4928	0.33
Admission age				**0.55**
1 day	4098	-0.0189	0.0254	0.46
2-7 days	855	0.0957	0.1079	0.38
8-14 days	118	-0.2837	0.2851	0.32
15-28 days	124	0.2354	0.2511	0.35
Place of Delivery				**<0.01**
KATH	3022	-0.3660	0.0381	<0.01
Referred	2173	0.5090	0.0530	<0.01
Discharge Diagnosis				**<0.01**
Prematurity	1285	-0.2410	0.1006	0.02
Respiratory Distress	288	1.1286	0.1460	<0.01
Infections	1222	-0.4296	0.0945	<0.01
Congenital anomalies	465	0.4858	0.1264	<0.01
Neonatal Jaundice	870	-0.6086	0.1199	<0.01
Birth asphyxia	945	0.8934	0.0925	<0.01
Others	120	-0.2591	0.3699	0.48

Table 3.9 Parameter estimates from the reduced model.

Variable	Sample size	Coefficient	Standard Error	p-value
Intercept		-2.7403	0.39354	<0.01
5 Minute Apgar				**<0.01**
0-3	202	1.8992	0.1683	<0.01
4-7	957	0.5034	0.0784	<0.01
8-10	4036	-0.2144	0.0209	< 0.01
Birth Weight				**<0.01**
VLBW	847	1.6005	0.1112	<0.01
LBW	1207	-0.1650	0.0792	0.03
Normal	3141	-0.3682	0.0452	<0.01
Gestational age				**<0.01**
Term	1050	-0.5601	0.1066	<0.01
Preterm	711	0.2472	0.1027	0.02
Not Stated	3434	0.1200	0.0330	<0.01
Discharge Diagnosis				**<0.01**
Prematurity	1285	-0.2505	0.1000	<0.01
Respiratory Distress	288	1.1191	0.1451	<0.01
Infections	1222	-0.4199	0.0935	<0.01
Congenital anomalies	465	0.4843	0.1260	<0.01
Neonatal Jaundice	870	-0.5842	0.1161	<0.01
Birth asphyxia	945	0.8746	0.0896	<0.01
Others	120	-0.2570	0.3694	0.49
Place of Delivery				**<0.01**
KATH	3022	-0.3780	0.0355	<0.01
Referred	2173	0.5256	0.0494	<0.01

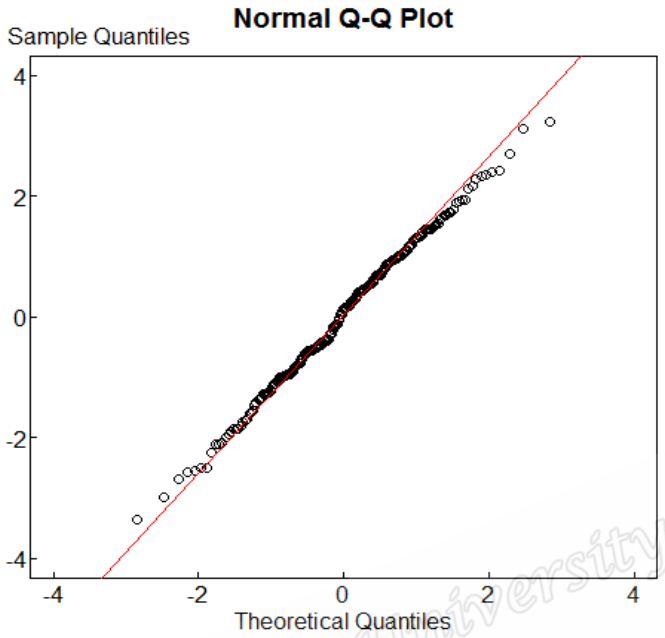

Figure 3.9a Normal quantile-quantile plot of residuals

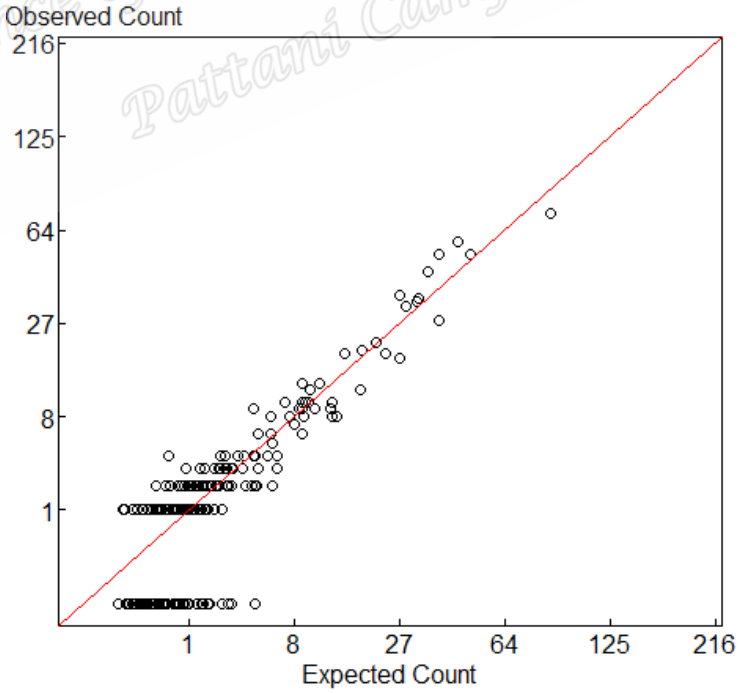

Figure 3.9b Expected and observed values

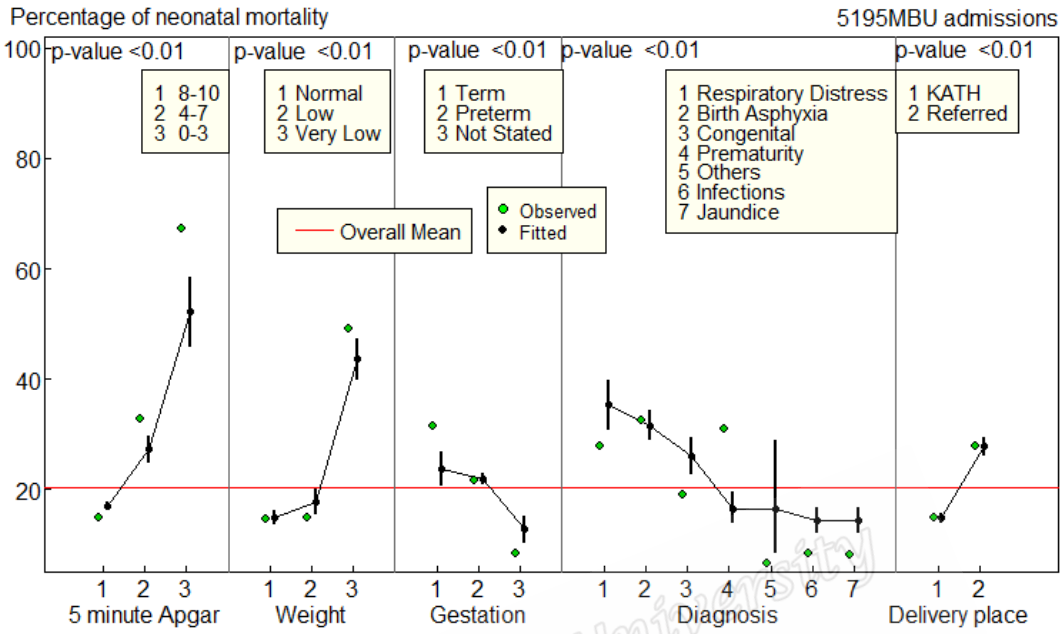

Figure 3.10 Observed and fitted probabilities from the logistic regression model

From the logistic regression analysis of the combined data, birth weight, 5 minute Apgar score, gestation, place of delivery and discharge diagnosis were significantly associated with neonatal death. Although gestational age was significant, its significance was only with respect to babies delivered at KATH. The analysis also showed no association of neonatal mortality with admission age and delivery mode. From the analyses, babies with 5 minute Apgar score less than 4 had very high percentage of mortality above the mean.

The ROC curve for the logistic regression model (Figure 3.11) showed that the model is able to predict 62.09% of mortality accurately. The five variables are able to

explain a large percentage of mortality. The independent variables explain different variations in the mortality.

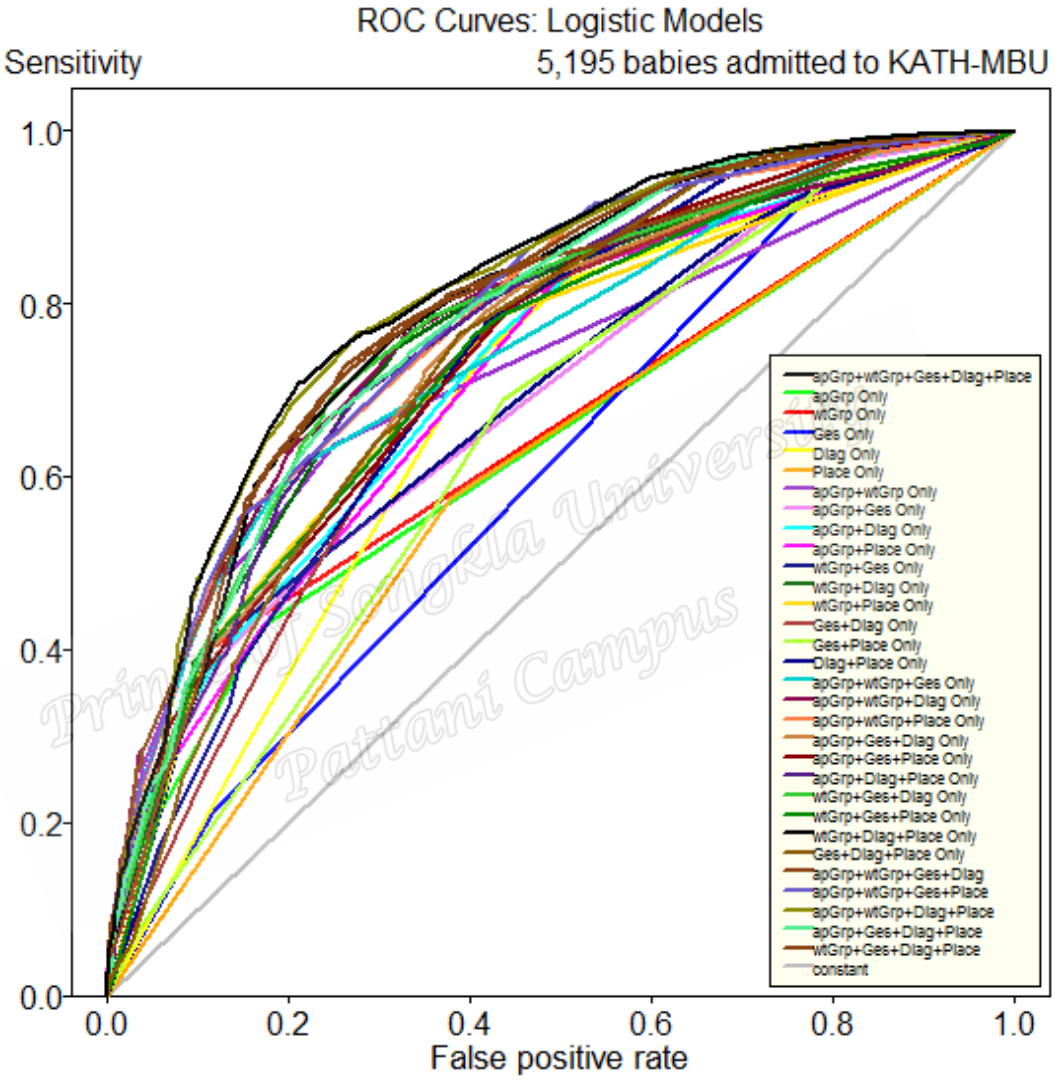

Figure 3.11 ROC curves for logistic regression model from total admissions

For instance, 5 minutes Apgar score alone explains 27.03%, birth weight explains 29.1%, gestation and discharge diagnosis explain 20.53% and 36.22% respectively, while place of delivery is responsible for 19.48% of the variations in mortality. Two variables,

Apgar score and birth weight are able to predict 44.23% of the neonatal mortality rate, birth weight and place of delivery predict 56.09% (Figure 3.12).

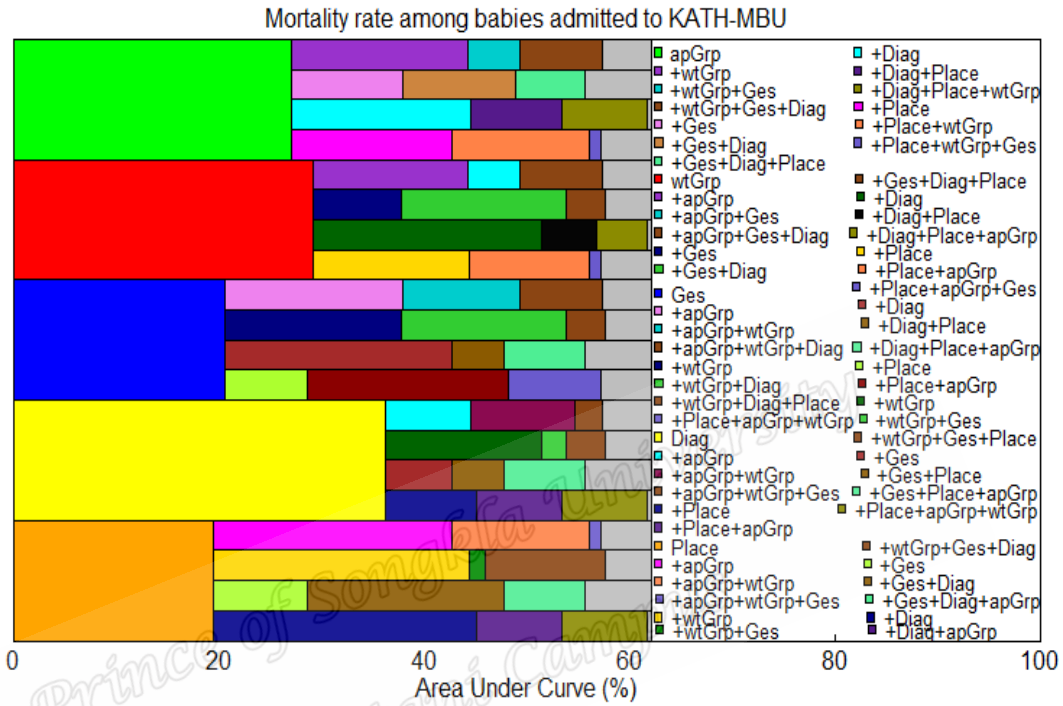

Figure 3.12 ROC bars for logistic regression model from total admissions

Chapter 4

Discussions and Conclusion

4.1 Summary of Results and Discussion

This study has shown that the likelihood function of the logistic regression model is regular. The likelihood function for all the models used in this thesis satisfied all the conditions for the existence and uniqueness of Maximum Likelihood Estimates (MLEs).

The logistic regression models have high predictive abilities. The model for babies referred to the neonatal unit has a predict power of 43.7%. This implies that the three variables are not enough to predict the mortality of most of the babies referred to the unit. The remaining 56.3% are largely due to other variables which may not have been considered in the model. Other demographic and maternal factors are likely to account for the remaining percentage.

The predictive ability of the logistic regression model for babies delivered at KATH is 72.53%. Thus, the variables investigated under this model are able to explain the causes of a large percentage of neonatal death among babies delivered at KATH. The smaller fraction-27.47% that is not explained by the model could be due to other variables that were not explored in the data. However, with such a high predictive ability, the logistic regression model could be said to be very accurate in explaining the neonatal mortality among these babies.

Findings from this analysis reveal that neonatal mortality in the neonatal unit is depended on the birth weight, five minute Apgar score, place of delivery, gestational age and discharge diagnosis. The proportions of death among babies referred to the neonatal unit are higher than that of babies delivered at KATH. Population and hospital based studies conducted in Pakistan and Ghana have also found that these factors are significant determinants of neonatal mortality (Jehan *et al.,* 2009; Welbeck *et al.,* 2003). There have been studies to show higher mortality rates in males than females (Jehan *et al.,* 2009; Seoud *et al.,* 2005), however, this study found otherwise. There was no significant difference between the fitted probabilities of males and females. However, the reasons for such contrast result could not be explained from the result. The neonatal mortality rate for babies delivered at KATH was 15 per 1000 live admissions while that of babies referred from other health facilities was 27.6 per 1000 live admissions. The overall mortality rate at the neonatal unit was 20.27 per 1000 live admissions. This result is lower than 32.1 and 38.7 per 1000 live births that was recorded for the KATH in 2008 and 2012 respectively (Siakwa *et al.,* 2014). The reduction could be attributed to improvements in resuscitation skills, equipment and technology at the neonatal unit and the increases in medical staff (Enweronu-Laryea *et al.*, 2008; Siakwa *et al.*, 2014). However, mortality rate among babies referred to the facility still remains very high. This is because most babies referred to KATH are usually very sick and require specialist care that is not provided by primary health facilities.

Although gestational age was a significant risk factor, it was only significant among babies delivered at KATH. Preterm birth was associated with a higher risk of death and

higher mortality rate (36.5 per 1000 live admissions) more than the overall mortality rate. This is comparable with results presented by Lawn et al., (2014) and Hsu et al., (2015). Preterm babies require special care from skilled workers to improve their chances of survival. Such special care involves providing warmth and feeding support (Lawn et al., 2013). The major conditions associated with neonatal death were infections, respiratory distress, prematurity, neonatal jaundice, sepsis, infections birth asphyxia and other neonatal conditions. Recent population based studies have also corroborated these findings (Edmond et al., 2008; GDHS, 2014; Welbeck et al., 2003). Among these neonatal conditions, birth asphyxia recorded the highest death rate. The mortality rate was 32.5 per 1000 live admissions. Clean birth practices have shown to significantly reduce neonatal mortality. Basic well know hygienic practices such as hand washing and maintaining a clean environment are poorly observed (Lawn et al., 2013).

In this present study birth weight was a major risk factor of neonatal death. The results revealed that VLBW had higher risk of neonatal mortality than babies with normal weight. Neonates with low birth weight and normal birth weight had mortality rates of 14.9 and 17.9 per 1000 live admissions respectively lower than overall mortality rate. Other studies in Sub-Saharan Africa (Adetola et al., 2011; Ajaari et al., 2012) and Asia (Arafa et al., 2003) have made similar findings to the effect that birth weight is a significant risk factor of neonatal mortality. However, the odds of death for low birth weight infants is lower than what was reported in a five-year study conducted at the facility in 2012 (Siakwa et al., 2014). This shows that the Neonatal Unit of the KATH has made great strides in saving low birth weight infants.

4.2 Conclusion

The neonatal mortality rate at the neonatal Unit of the KATH is still high. This study found that babies referred from home or other health facilities, babies having very low birth weight, 5 minute Apgar score of less than 4, preterm birth and having congenital anomalies were more likely to have a higher neonatal mortality rate. There is, therefore, the need for continuous attention and strengthening of newborn interventions to help reduce the risk reduce the risk of mortality among neonates delivered at KATH

4.3 Study limitations

This was a hospital based study and not a representation of the entire population in the city. Data for this study is retrospective and thus, there were cases of missing information although not significant. The parameter estimates from the logistic regression model were only verified for normality. Therefore, a further study is needed to explore the large sample behavior of the estimates.

Another limitation to this study is with regards to the logistic regression model for babies referred to the neonatal unit from different health facilities. The three variables investigated for the data are able to explain only 43.7% of the variations in mortality. 5 minute Apgar score explains an insignificant amount of the variations. Hence, it is suggested that high level of information is collected from the babies referred to the unit in order to explore other variables that may explain the remaining 57% of the variations.

References

Abreu, M.N.S., Siqueira, A.L. and Caiaffa, W.T. 2009. Ordinal logistic regression in epidemiological studies. Revista de Saude Publica. 43(1), 183-194.

Adetola, A.O., Tongo, O.O., Orimadegun, A.E. and Osinusi, K. 2011. Neonatal mortality in an urban population in Ibadan, Nigeria. Pediatrics and Neonatology. 52(5), 243-250.

Ajaari, J., Masanja, H. and Weiner, R. 2012. Impact of Place of Delivery on Neonatal Mortality in Rural Tanzania. International Journal of MCH and AIDS. 1(1), 49-59.

Albert, A., and Anderson, J.A. 1984. On the existence of maximum likelihood estimates in logistic regression. Biometrika, 71(1), 1-10.

Aldrich, J. 1997. R. A. Fisher and the Making of Maximum Likelihood 1912–1922. Statistical Science, 12, 162-176

Antia, H.M. 2002. Numerical Methods for Scientists and Engineers. Springer Verlag New York Inc., New York, U.S.A., pp. 302.

Arafa, M. and Alshehri, M. 2003. Predictors of neonatal mortality in the intensive care unit in Abha, Saudi Arabia. Saudi Medical Journal. 24(12), 1374-1376.

Bacak, S.J., Baptiste-Roberts, K., Amon, E., Ireland, B. and Leet, T. 2005. Risk factors for neonatal mortality among extremely-low-birth-weight infants. American Journal of Obstetrics and Gynecology. 192, 862-867.

Bewick, V., Cheek, L. and Ball, J. 2005. Statistics review 14: Logistic regression. Critical Care London, England. 9(1), 112-118.

Bhutta, Z A., Das, J.K., Bahl, R., Lawn, J.E., Salam, R.A., Paul, V K. and Walker, N. 2014.

Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? The Lancet. 384(9940), 347-70.

Co, T. B. 2013. Methods of Applied Mathematics for Engineers and Scientists. Cambridge University Press.New York, U. S. A., pp 37-38

Czepiel, S.A. 2002. Maximum likelihood estimation of logistic regression models: theory and implementation. Available at czep. net/stat/mlelr.

Dickson, K.E., Simen-Kapeu, A., Kinney, M.V, Huicho, L., Vesel, L., Lackritz, E. and Lawn, J.E. 2014. Every Newborn: health-systems bottlenecks and strategies to accelerate scale-up in countries. The Lancet. 384(9941), 438-454.

Edmond, K.M. 2006. Delayed Breastfeeding Initiation Increases Risk of Neonatal Mortality. Pediatrics. 117(3), e380-e386.

Edwards, A. W. F. 1984. Likelihood. Cambridge University Press Archive. pp 145.

Enweronu-Laryea, C.C., Nkyekyer, K. and Rodrigues, O.P. 2008. The impact of improved neonatal intensive care facilities on referral pattern and outcome at a teaching hospital in Ghana. Journal of Perinatology. 28(8), 561-565.

Fox, J. 2008. Applied regression analysis and generalized linear models. 2nd ed.,. Sage Publications Inc., New York, U.S.A., pp 379-380.

Garthwaite, P. H., Jolliffe, I. T., & Jones, B. 2002. Statistical inference. Oxford University Press on Demand.pp45

Givens, G.H. and Hoeting, J.A. 2005. Computational Statistics, John Wiley & Sons, Inc.

Golub, G.H. and van Loan, C.F. 1996. Matrix computations 3rd ed., John Hopkins University Press, London, UK., pp 140.

Goodman, D.C., Fisher, E.S., Little, G.A., Stukel, T.A., Chang, C.H. and Schoendorf, K.S. 2002. The relation between the availability of neonatal intensive care and neonatal mortality. New England Journal of Medicine. 346(20), 1538-1544.

Gourieroux, C. and Monfort, A. 1981. Asymptotic properties of the maximum likelihood estimator in dichotomous logit models. Journal of Econometrics, 17(1), 83-97.

Greene, W.H. 2005. Econometric Analysis 5th ed. Prentice Hall, New Jersey, U.S.A., pp. 473.

Greene, W.H. 2005. Econometric Analysis 5th ed. Prentice Hall, New Jersey, U.S.A., pp. 479.

GSS, G, and Macro, I.C.F. 2014. Ghana demographic and health survey 2013. Accra, Ghana. Ghana Statistical Service, Ghana Health Service, and ICF Macro.

GSS, G. and Macro, I.C.F. 2009. Ghana demographic and health survey 2008. Accra, Ghana. Ghana Statistical Service, Ghana Health Service, and ICF Macro.

Gustafson, K.E. 1999. Introduction to Partial Differential Equations and Hilbert Space Methods. 3rd ed. Dover Publications Inc, New York, U.S.A., pp 126.

Hosmer Jr, D.W. and Lemeshow, S. 2004. Applied logistic regression. John Wiley & Sons Inc., New York, U.S.A., pp. 95.

Hsu, S.T., Hsieh, C.J., Chen, H.W., Jeng, S.F., Wu, H.C., Chen, M.H. and Hsieh, W.S. 2014. Nationwide Birth Weight and Gestational Age-specific Neonatal Mortality Rate in Taiwan. Pediatrics and Neonatology. 56(3), 149-158.

Kayode, G.A., Ansah, E., Agyepong, I., Amoakoh-Coleman, M., Grobbee, D.E. and Klipstein-Grobusch, K. 2014. Individual and community determinants of neonatal

mortality in Ghana: a multilevel analysis. BMC Pregnancy and Childbirth. 14(1), 165.

Kazembe, L.N. and Kandala, N.B. 2015. Estimating areas of common risk in low birth weight and infant mortality in Namibia: A joint spatial analysis at sub-regional level. Spatial and Spatio-Temporal Epidemiology. 12, 27-37.

Kinney, M.V., Cocoman, O., Dickson, K.E., Daelmans, B., Zaka, N., Rhoda, N.R. and Darmstadt, G.L. 2015. Implementation of the Every Newborn Action Plan: Progress and lessons learned. Seminars in Perinatology. 39(5), 326-337.

Kleinbaum, D.G., Kupper, L.L., Muller, K.E., and Nizam, A. 1998. Applied Regression Analysis and Other Multivariable Methods. Duxbury Press, New York, U.S.A., pp. 317.

Klingenberg, C., Olomi, R., Oneko, M., Sam, N. and Langeland, N. 2003. Neonatal morbidity and mortality in a Tanzanian tertiary care referral hospital. Annals of Tropical Paediatrics. 23, 293-299.

Lander, T. 2006. Neonatal and Perinatal mortality: country, regional and global estimates. World Health Organization.

Lawn, J.E., Blencowe, H., Oza, S., You, D., Lee, A.C., Waiswa, P. and Cousens, S.N. 2014. Every Newborn: progress, priorities, and potential beyond survival. The Lancet. 384(9938), 189-205.

Lourakis, M.I.A., and Argyro, A.A. 2005. Is Levenberg-Marquardt the Most Efficient Optimization Algorithm for Implementing Bundle Adjustment?. In: ICCV, pp. 1526–1531 Available at http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1544898&tag=1

Mai, T., Bastin, F., Toulouse, M., 2014. On Optimization Algorithms for the Maximum Likelihood Estimation. CIRRELT Technical Report, No. 64.

McNeil, D. 2016. Workshop on Research Methods for Ecologists http://www.st.sat.psu.ac.th/workshop/CMU16032016/index.html

Moré, J.J. and Cosnard, M.Y. 1979. Numerical solution of nonlinear equations. ACM Transactions on Mathematical Software (TOMS). 5(1), 64-85.

Myung, I. J. 2003. Tutorial on maximum likelihood estimation. Journal of Mathematical. Psychology. 47, 90-100

Nielsen, F. (2013). Cramer-Rao lower bound and information geometry.arXiv preprint arXiv:1301.3578.

Nisar, Y.B and Dibley, M.J. 2014. Determinants of neonatal mortality in Pakistan: secondary analysis of Pakistan Demographic and Health Survey 2006-07. BMC Public Health. 14(1), 663.

Okeh, U.M, and Oyeka I.C.A. 2013. Estimating the Fisher's Scoring Matrix Formula from Logistic Model. American Journal of Theoretical and Applied Statistics 2(6):221-227.

Pawitan, Y. 2001. In all likelihood: Statistical Modelling and Inference using likelihoods. Oxford University Press. United Kingdom, pp. 241.

Pawitan, Y. 2001. In all likelihood: Statistical Modelling and Inference using likelihoods. Oxford University Press. United Kingdom, pp. 231.

Qi, F. and Chen, S. 2007. Complete monotonicity of the logarithmic mean. Mathematical Inequalities and Applications. 10(4), 799-804.

Sankaran, K., Chien, L.Y., Walker, R., Seshia, M., Ohlsson, A. and Lee, S.K. 2002. Variations in mortality rates among Canadian neonatal intensive care units. Canadian Medical Association Journal. 166(2), 173-178.

Shiffman, J. 2015. Network advocacy and the emergence of global attention to newborn survival, Oxford Univ Press. 30(12), 1-13.

Smyth, G.K. 1998. Optimization and nonlinear equations. In: Armitage,P. and Colton,T. (eds), Encyclopedia of Biostatistics.. Wiley, London, pp. 3174–3180.

Taylor, S., Williams, B., Magnus, D., Goenka, A. and Modi, N. 2015. From MDG to SDG: good news for global child health? The Lancet. 386(10000), 1213-1214.

Tongkumchum, P., and McNeil D. 2009. Confidence Intervals Using Contrasts for Regression Model. Songklanakarin journal of Science and Technology 31(2): 151–56.

UNICEF. 2015. Committing to Child Survival: a Promise Renewed-Progress Report. United Nations Children's Fund,The (UNICEF).

Van den Bos, A. 2005. Parameters Estimation for Scientiests and Engineers. John Wiley and Sons Inc. New Jersey, U.S.A., pp 201.

van der Vaart, A.W. 1998. Asymptotic Statistics. Cambridge: Cambridge University Press

Welbeck, J., Biritwum, R. B., and Mensah, G. 2003. Factors affecting the survival of the “at risk” newborn at Korle Bu Teaching Hospital, Accra, Ghana. West African journal of medicine, 22(1), 55-58.

White, G.C. 2007. The Binomial Likelihood Function. 8-10. Available at http://warnercnr.colostate.edu/~gwhite/fw663/BinomialLikelihood.pdf

Yee, T. W. 2015. Vector Generalized Linear and Additive Models: With an Implementation in R. Springer Series in Statistics. New York, U.S.A., pp541.

You, D., Hug, L., Ejdemyr, S., Beise, J., Idele, P., Gerland, P. and Spoorenberg, T. 2015. Child Mortality. World Health Organization.

Appendix

United Nations
Educational, Scientific and
Cultural Organisation

In Support of
UNESCO's 70th
Anniversary Celebrations

Choose to be Huge
International
Multidisciplinary
Research Foundation

INTERNATIONAL MULTIDISCIPLINARY ACADEMIC CONFERENCE THAILAND 2015

T022A

Certificate of Participation & Presentation

This is to certify that

MR. OWUSU BENJAMIN ATTA

has participated in the
International Multidisciplinary Academic Conference Thailand 2015
and presented a paper titled

NEONATAL MORTALITY AMONG INFANTS REFERRED TO THE MOTHER AND BABY UNIT, KOMFO ANOKYE TEACHING HOSPITAL, KUMASI, GHANA

held at IMRF Thailand Chapter, Jomtien, Pattaya, Thailand,
Nov 02-06, 2015.

Dr. Jaganmohan Bajaj
UNESCO Expert,
Special Envoy, Ministry of Education,
Govt of Republic of Kosovo

Dr. Ratnakar D Bala
Director (Academics)
International Multidisciplinary
Research Foundation

Neonatal mortality among infants referred to the Teaching Hospital, Kumasi, Ghana

Owusu Benjamin Atta[1], Apiradee Lim[1] Priscilla Wobil[2]

[1] Department of Mathematics and Computer Science, Prince of Songkla University, Pattani Campus, Pattani 94000, Thailand.

[2] Child Health Directorate, Komfo Anokye Teaching Hospital, Kumasi, Ghana.

Abstract

Background: The first 28 days of life- the neonatal period is the most vulnerable time for a child's survival. Globally, neonatal mortality has seen a downward trend in recent years. Understanding the risk factors associated with neonatal mortality at the neonatal unit is important because it allows inferences about the quality of care.

Objective: To determine the neonatal mortality rate of infants referred to the Neonatal Unit of the Komfo Anokye Teaching Hospital, Kumasi, Ghana.

Methods: Data were obtained from the treatment files of neonates admitted at the Neonatal Unit during the period from January 2013 through to May 2014. In all, a total of 2173 neonates were used for this study. The babies were grouped by weight, delivery mode and admission age. Neonatal mortality was defined as death that occurred within the first 28 days of life. Logistic regression models were fitted to the data to assess the association between neonatal and these predictors while adjusting for confounders.

Results: The mortality rate among VLBW is almost 40% higher than normal birth weight. Vaginally delivered babies had a 2.5% higher rate than those delivered through caesarean section and the infants that were admitted within the first 24 hours of birth had a mortality rate 13% higher than those admitted after 24 hours. The neonatal mortality rate amongst infants referred to the facility was 27.7 per 1000 live births.

Keywords: Neonatal mortality, Kumasi.

Background

The first four weeks of life are the most crucial period when infants are highly prone to illness and death (1). Each year, almost 3 million babies die within the neonatal period (2). Latest reports indicate that over 5.9 million under-five deaths were recorded in 2015 with daily average of 16 000 deaths (3). Globally, newborn deaths accounts for almost 44% of under-five deaths with annual estimates of almost 2.8 million (4). Majority of these neonatal deaths occur in developing countries. In fact, neonatal mortality accounts for almost half of all infant deaths in these countries (5). Preventive measures and interventions have helped to reduce neonatal mortality from 36 per 1000 live births in 1990 to 19 per 1000 live births in 2015 (6). Globally, neonatal mortality has been decreasing at a slower rate from 1990 as compared to under-five mortality (7). The decline in neonatal death especially in Ghana from 32 per 1000 live births to 29 per 1000 live births has partly been attributed to improved facilities at the neonatal units and improved neonatal care (8). However, neonatal mortality still remains a major challenge for Ghana and other lower middle income countries particularly in Sub-Saharan Africa and southern Asia; Countries in these two regions have made the least progress towards reducing neonatal mortality (9). The year 2015 is fast running to a close and less than 100 countries have been able to achieved Millennium Development Goal-4 of reducing under-five mortality by two-thirds (3). Through the newly formulated Sustainable Development Goals, UN member nations have committed themselves to reducing under-five mortality and neonatal mortality to 25 per 100 live birth and 15 per 1000 live births respectively (10). Infant and neonatal mortality rates are significant measures of health quality, societal welfare and socio-economic status of a country (11-13). A detailed study into the predictors of neonatal mortality at the various neonatal units is necessary as it allows inferences about the quality of care. Predictors of neonatal mortality give insights into how neonates could be managed to improve the outcomes of admissions at the neonatal unit. Separate studies have identified place of birth, birth weight, mode of delivery, delayed breastfeeding initiation and age of baby at the time of admission as significant predictors of neonatal death (12, 14-16). This study aimed to investigate the prevalence of neonatal mortality amongst infants delivered

and admitted to the Mother and Baby Unit, Komfo Anokye Teaching Hospital, Kumasi, Ghana.

Methods

This was a hospital based study which used retrospective primary data from the neonatal unit to identify the major predictors of neonatal death at the hospital. This study was conducted at the neonatal unit-Mother and Baby Unit (MBU) of the Komfo Anokye Teaching Hospital from January 2013 to May 2014. The Komfo Anokye Teaching Hospital (KATH) is a tertiary health facility located in a Kumasi, the capital of the Ashanti region and the most second most populous city in Ghana. KATH serves as the main referral facility for the Northern Sector of the country. The Neonatal unit has three wards. There is a High Dependency Unit that admits sick babies referred from the delivery wards. There is also a Preterm/Low birth weight (LBW)/Kangaroo Mother Care unit that admits preterm/LBW babies from the High Dependency unit who have been stabilized and there is a Septic unit that admits out-born sick newborns and infants up to the age of two months. There are incubators and phototherapy machines in the unit. Different categories of babies are admitted to the facility. These include preterm babies, low birth weight babies, babies with neonatal jaundice and sepsis, babies with congenital anomalies.

All neonates admitted to the unit during the period from January 2013 through to May 2014 were included in the study. They were followed until discharged or death. The following information was retrieved from the in-patient files: birth weight, sex, mode of delivery, place of delivery, age on admission, Apgar scores, time between birth and admission and outcome of admission. In all, a total of 2173 neonates were used for this study. The babies were grouped by weight, delivery mode, and admission age. Neonatal mortality was defined as death that occurred within the first 28 days of life. The outcome of this study was neonatal mortality.

Birth weights, mode of delivery and babies' age at the time of admission were employed as exposure variables in this study. The birth weight was categorized into three groups namely; very low birth weight (VLBW), low birth weight (LBW) and normal weight. VLBW was defined as weights less than 1.5kg; LBW for those between 1.5kg-2.4kg,

normal birth weight for those above 2.5kg. The mode of delivery was divided into two categories. The first group was vaginal delivery which included all neonates delivered by spontaneous delivery and assisted delivery (vacuum extraction). Babies delivered through caesarean section were also categorized as one. The age as of the time of admission was divided into two groups; either the baby was less than 24 hours old or more than 24 hours.

Statistical analysis

Descriptive statistics was used to calculate the neonatal mortality rate at the Neonatal Unit and measures of associations between the exposure variables. Univariate analysis was done to examine the variables associated with neonatal death. These variables which were significant in the univariate analysis were consequently included in the multiple logistic regression models. Multiple logistic regression models were used to determine the strength of association between these predictors and the outcome while adjusting for confounders.

Ethical approval

Ethical consent was obtained from Research and Development Unit of KATH and the Committee on Human Research, Publications and Ethics, Kwame Nkrumah University of Science and Technology, School of Medical Sciences and Komfo Anokye Teaching Hospital, Kumasi.

Results:

Factors such as birth weight, delivery mode, admission age of the baby were all found to be significantly associated with neonatal death.

Neonates who had VLBW (OR = 47.8; 95% CI = (29.24-77.99)) had more than forty times higher risk of death compared to neonates who had normal birth weight. VLBW babies also had seven times higher risk of death (OR= 7.3; 95% CI = (5.35-10.06)). Babies who were delivered through Vaginally (OR = 1.62; 95% CI = (1.28, 2.05)) had a slightly higher risk of death in comparison to babies delivered through Caesarean section.

The risk of death among neonates admitted within the first 24 hours of birth was almost three times higher (OR = 2.94; CI = (1.43, 6.06)). The neonatal mortality rate amongst

infants referred to the facility was 27.7 per 1000 live births and this is indicated by the horizontal read line (Fig. 1)

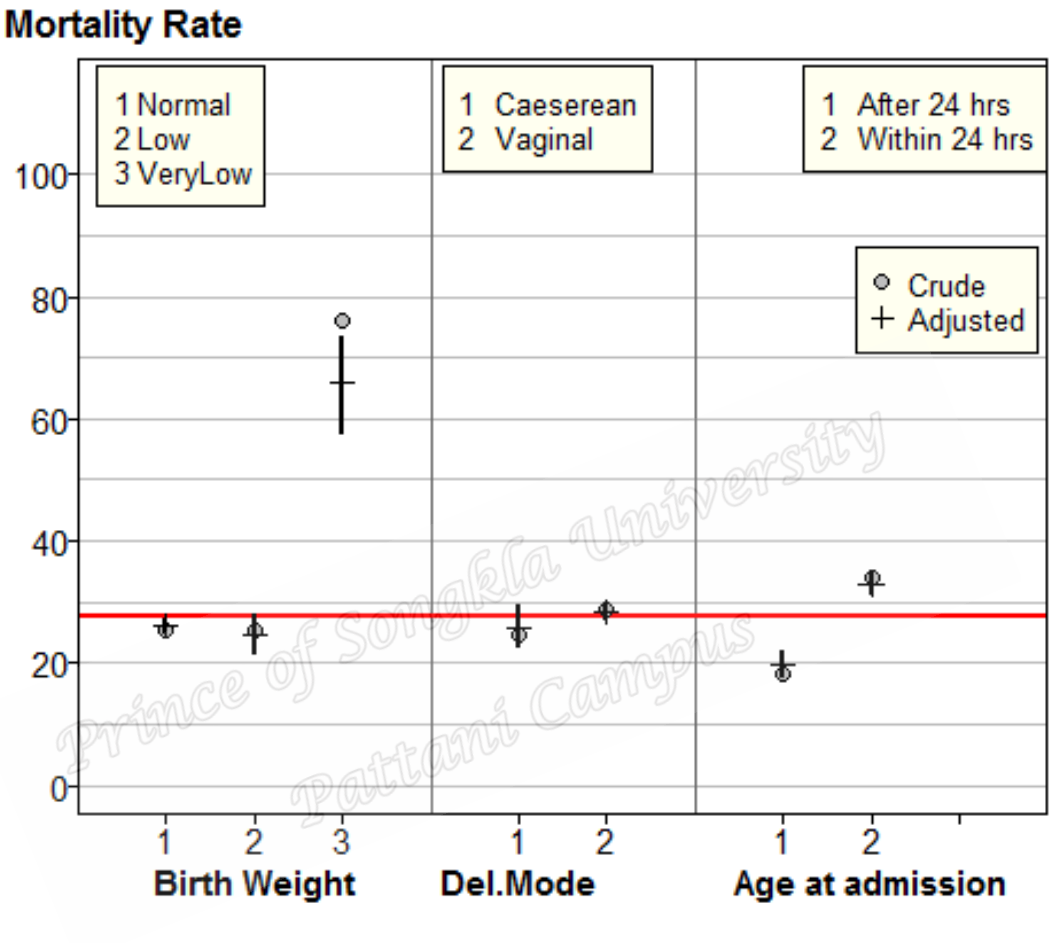

Fig. 1: Mortality rate of each variable used in the logistic regression model

Discussion

This was a hospital based study and not a representation of the entire population in the city. Our study revealed that the survival of a neonate in the Neonatal Unit depends on the birth weight, delivery mode and the admission age. These findings are consistent with similar studies that have been conducted on neonatal mortality (17, 18). Although different researches have documented evidence to show higher odds of death in males than females (19, 20), our multivariate analysis revealed otherwise. We found no significant difference in the risks of death and mortality rates between males and females just like studies that

were conducted in Ibadan (1), Tanzania (5), Saudi Arabia (11) and Cairo (21). The overall mortality rate at the unit is 14.9 per 1000 live births. This indicates a drastic reduction from the figures that were reported for the same facility between 2008 and 2012 (21). However, the neonatal mortality rate reported by the Ghana Demographic Health Survey for the Ashanti region and other studies conducted in other parts of Ghana is higher than our finding for the study period (22-24). The difference could be attributed to the fact that these studies were population based studies and considered larger sample sizes.

The logistic regression model fitted the data quite well and revealed that ELBW had the highest risk of neonatal mortality at the unit and babies in this category of weight are the most vulnerable. VLBW babies had over forty fold risk of death compared with normal birth weight babies and a mortality rate of 57.8 per 1000 live births (Fig 1) which is far above the mean mortality rate. The risk of death of LBW babies was almost the same as normal birth weight babies and there was little difference in their mortality rates. Neonates with low birth weight and normal birth weight had mortality rates of 11.8 and 11.2 per 1000 live births respectively lower than overall mortality rate; while big babies had a mortality rate of 7.8 per 1000 lives. Other studies in Sub-Saharan Africa (1, 14) and Asia (11, 26) have made similar findings to the effect that birth weight is a significant risk factor of neonatal mortality. However, the odds of death for low birth weight infants is lower than what was reported in a five-year study conducted at the facility in 2012 (22). This shows that the Neonatal Unit of the KATH has made great strides in saving low birth weight infants.

Kangaroo Mother Care (KMC) has so far proven to be very effective in improving survival chances among preterm, very low and low birth weight infants. Continuous implementation of the KMC program will enhance the survival chances of such babies (27, 28)

Improving the facilities in the neonatal unit as well as improving resuscitation skills also has a positive effect in saving normal weight but asphyxiated and jaundiced babies (1, 8, 29-31).

In the present study, we found an association between mode of delivery and neonatal death. The risk of neonatal death was higher among babies delivered by Vaginal with (OR= 1.6; CI = (1.28, 2.05)) compared to babies delivered through caesarean section and they have higher mortality rate as well. Other studies have documented similar findings to prove that there is high risk of death among very low birth weight and preterm infants delivered spontaneously and that it was much safer to deliver preterm babies by caesarean delivery especially in breech presentation (32-37). However, opinions are still varied on the significance of mode of delivery as a risk factor of neonatal death as another study found no significance association between the two (38). The risk of neonatal death may not necessarily be caused by the mode of birth but physician's reluctance to intervene for the fetus with the perception that the infant is too young or small to survive ((35, 36, 39).

A high proportion of the babies were admitted within the first 24 hours of birth and their mortality rate was above the mean percentage of death (Fig. 1). The associated risk is almost three folds compared to those who were admitted after 24 hours. This indicates how fragile these babies are within the first day. Some of these admissions are caused by neonatal infections. Lack of sufficient antenatal care is highly correlated with these neonatal infections (40, 41)

Conclusion

The neonatal mortality rate at the neonatal Unit of the KATH is still high. This study found that babies with extremely low birth weight, very low birth weight, babies delivered by spontaneous vaginal delivery and babies less 24 hours are at the highest risk of death. There is the need for urgent attention and interventions to help reduced the risk associated with these neonates. Such interventions include improving the facilities at the neonatal unit and improving the resuscitation skills of health workers.

Funding

This work was supported by the Higher Education Research Promotion and the Thailand's Education Hub for Southern Region of ASEAN Countries Project Office of the higher Education Commission.

Acknowledgement

Our tremendous thanks go to Emeritus Prof. Dr. Donald McNeil for his comments and support during this study and to the Mother and Baby Unit, KATH for providing the data on neonatal admissions.

References

1. Adetola AO, Tongo OO, Orimadegun AE, Osinusi K. Neonatal mortality in an urban population in Ibadan, Nigeria. Pediatrics & Neonatology. 2011;52(5):243-50.
2. Shiffman J. Network advocacy and the emergence of global attention to newborn survival. Health policy and planning. 2015:czv092.
3. UNICEF. Committing to Child Survival: A Promise Renewed, Progress Report 2015, UNICEF. New York. 2015.
4. Kinney MV, Cocoman O, Dickson KE, Daelmans B, Zaka N, Rhoda NR, et al., editors. Implementation of the Every Newborn Action Plan: progress and lessons learned. Seminars in Perinatology; 2015: Elsevier
5. Klingenberg C, Olomi R, Oneko M, Sam N, Langeland N. Neonatal morbidity and mortality in a Tanzanian tertiary care referral hospital. Annals of Tropical Paediatrics: International Child Health. 2003;23(4):293-9
6. UNICEF. UN Inter Agency Group for Child Mortality Estimation 2015. Levels and Trends in Child Mortality. 2015.
7. Lawn JE, Blencowe H, Oza S, You D, Lee AC, Waiswa P, et al. Every Newborn: progress, priorities, and potential beyond survival. The Lancet. 2014;384(9938):189-205.
8. Enweronu-Laryea C, Nkyekyer K, Rodrigues O. The impact of improved neonatal intensive care facilities on referral pattern and outcome at a teaching hospital in Ghana. Journal of Perinatology. 2008;28(8):561-5.
9. Dickson KE, Simen-Kapeu A, Kinney MV, Huicho L, Vesel L, Lackritz E, et al. Every Newborn: health-systems bottlenecks and strategies to accelerate scale-up in countries. The Lancet. 2014;384(9941):438-54.

10. Taylor S, Williams B, Magnus D, Goenka A, Modi N. From MDG to SDG: good news for global child health? The Lancet. 2015;386(10000):1213-4
11. Arafa MA, Alshehri MA. Predictors of neonatal mortality in the intensive care unit in Abha, Saudi Arabia. Saudi medical journal. 2003;24(12):1374-6.
12. Hsu S-T, Hsieh C-J, Chen H-W, Jeng S-F, Wu H-C, Chen M-H, et al. Nationwide Birth Weight and Gestational Age-specific Neonatal Mortality Rate in Taiwan. Pediatrics & Neonatology. 2014.
13. Kazembe LN, Kandala N-B. Estimating areas of common risk in low birth weight and infant mortality in Namibia: A joint spatial analysis at sub-regional level. Spatial and spatio-temporal epidemiology. 2015;12:27-37.
14. Ajaari J, Masanja H, Weiner R, Abokyi SA, Owusu-Agyei S. Impact of place of delivery on neonatal mortality in rural Tanzania. International Journal. 2012;1(1):49-59.
15. Brown HK, Speechley KN, Macnab J, Natale R, Campbell MK. Neonatal morbidity associated with late preterm and early term birth: the roles of gestational age and biological determinants of preterm birth. International journal of epidemiology. 2013:dyt251.
16. Edmond KM, Zandoh C, Quigley MA, Amenga-Etego S, Owusu-Agyei S, Kirkwood BR. Delayed breastfeeding initiation increases risk of neonatal mortality. Pediatrics. 2006;117(3):e380-e6.
17. Bacak SJ, Baptiste-Roberts K, Amon E, Ireland B, Leet T. Risk factors for neonatal mortality among extremely-low-birth-weight infants. American journal of obstetrics and gynecology. 2005;192(3):862-7.
18. Jehan I, Harris H, Salat S, Zeb A, Mobeen N, Pasha O, et al. Neonatal mortality, risk factors and causes: a prospective population-based cohort study in urban Pakistan. Bulletin of the world Health Organization. 2009;87(2):130-8.
19. Okolo A, Omene J. Trends in neonatal mortality in Benin City, Nigeria. International Journal of Gynecology & Obstetrics. 1985;23(3):191-5.

20. Onayade A, Sule S, Elusiyan J. Determinants of neonatal mortality at Wesley Guild Hospital, Ilesa, Nigeria. Nigerian journal of medicine: journal of the National Association of Resident Doctors of Nigeria. 2005;15(3):271-6.
21. Seoud I, El-Din RMG, Said RN, Hessin HA. Predictors of Neonatal Mortality in Intensive Care Unit in Children's Hospital, Cairo University. Alexandria Journal of Pediatrics. 2005;19(1):93.
22. Siakwa M, Kpikpitse D, Laryea T, Ankobil A, Dare S, Ebu N. A five-year neonatal mortality trend in a Ghanaian Teaching Hospital after the implementation of strategies to achieve the Millennium Development Goal (MDG) 4. International Journal of Pediatrics and Child Health. 2014;2(1):043-049.
23. GSS G, Macro I. Ghana demographic and health survey 2008. Accra, Ghana: Ghana Statistical Service, Ghana Health Service, and ICF Macro. 2009.
24. Welaga P, Moyer CA, Aborigo R, Adongo P, Williams J, Hodgson A, et al. Why are babies dying in the first month after birth? A 7-year study of neonatal mortality in northern Ghana. PloS one. 2013;8(3):e58924.
25. Engmann C, Walega P, Aborigo RA, Adongo P, Moyer CA, Lavasani L, et al. Stillbirths and early neonatal mortality in rural Northern Ghana. Tropical Medicine & International Health. 2012;17(3):272-82.
26. Nisar YB, Dibley MJ. Determinants of neonatal mortality in Pakistan: secondary analysis of Pakistan Demographic and Health Survey 2006-07. BMC public health. 2014;14(1):663.
27. Lawn JE, Mwansa-Kambafwile J, Horta BL, Barros FC, Cousens S. 'Kangaroo mother care'to prevent neonatal deaths due to preterm birth complications. International journal of epidemiology. 2010;39(suppl 1):i144-i54.
28. Charpak N, Ruiz-Peláez JG, Zita Figueroa de C M, Charpak Y. Kangaroo mother versus traditional care for newborn infants≤ 2000 grams: a randomized, controlled trial. Pediatrics. 1997;100(4):682-8.

29. Carlo WA, Wright LL, Chomba E, McClure EM, Carlo ME, Bann CM, et al. Educational impact of the neonatal resuscitation program in low-risk delivery centers in a developing country. The Journal of pediatrics. 2009;154(4):504-8. e5.
30. Bookman L, Engmann C, Srofenyoh E, Enweronu-Laryea C, Owen M, Randolph G, et al. Educational impact of a hospital-based neonatal resuscitation program in Ghana. Resuscitation. 2010;81(9):1180-2.
31. Enweronu-Laryea C, Engmann C, Osafo A, Bose C. Evaluating the effectiveness of a strategy for teaching neonatal resuscitation in West Africa. Resuscitation. 2009;80(11):1308-11.
32. Barzilay E, Mazaki-Tovi S, Amikam U, de Castro H, Haas J, Mazkereth R, et al. Mode of delivery of twin gestation with very low birthweight: is vaginal delivery safe? American journal of obstetrics and gynecology. 2015.
33. Herbst A, Källén K. Influence of mode of delivery on neonatal mortality and morbidity in spontaneous preterm breech delivery. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2007;133(1):25-9.
34. Shamsa A, Bai J, Raviraj P, Gyaneshwar R. Mode of delivery and its associated maternal and neonatal outcomes. Open Journal of Obstetrics and Gynecology. 2013;03(03):307-12.
35. Jonas HA, Khalid N, Schwartz SM. The relationship between Caesarean section and neonatal mortlity in very-low-birthweight infants born in Washington State, USA. Paediatric and perinatal epidemiology. 1999; 13(2):170-189.
36. Redman ME, Gonik B. Cesarean delivery rates at the threshold of viability. American Journal of Obstetrics and Gynecology. 2002; 187(4):873-6.
37. Armson BA, O'Connell C, Persad V, Joseph KS, Young DC, Baskett TF. Determinants of perinatal mortality and serious neonatal morbidity in the second twin. Obstetrics & Gynecology. 2006; 108(3):556-564.
38. Alburke S, Ashur B, Assadi M. Neonatal and Perinatal Mortality Rates in Neonatal Unit of Misurata Teaching Hospital–Libya/2013. J Hematol Thrombo Dis 3: 184. doi:

10.4172/2329-8790.1000184 Page 2 of 6 J Hematol Thrombo Dis ISSN: 2329-8790 JHTD, an open access journal Volume 3• Issue 1• 1000184. Total. 2015;84(100):3.

39. Bottoms S, Paul R, Iams J, Mercer B, Thom E, Roberts J, et al. Obstetric determinants of neonatal survival: influence of willingness to perform cesarean delivery on survival of extremely low-birth-weight infants. American journal of obstetrics and gynecology. 1997;176(5):960-6.
40. Mizumoto BR, Moreira BM, Santoro-Lopes G, Cunha AJ, Santos RM, Pessoa-Silva CL, et al. Quality of antenatal care as a risk factor for early onset neonatal infections in Rio de Janeiro, Brazil. Braz J Infect Dis. 2015;19(3):272-7.
41. Mizumoto BR, Moreira BM, Santoro-Lopes G, Cunha AJ, Santos RM, Pessoa-Silva CL, et al. Quality of antenatal care as a risk factor for early onset neonatal infections in Rio de Janeiro, Brazil. Braz J Infect Dis. 2015;19(3):272-7.

Vitae

Name: Benjamin Atta Owusu

Student ID: 5720320702

Educational Attainment:

Degree	Name of Institution	Year of Graduation
B.Sc. (Mathematics)	Kwame Nkrumah University of Science and Technology	2013

Awards and Honors:

Thailand's Education Hub for Southern Region of ASEAN Countries (TEH-AC) Scholarship.

Proceedings:

Owusu, .B., A., Lim, A., and Wobil, P. 2015. Neonatal mortality among infants referred to the Teaching Hospital, Kumasi, Ghana. International Multidisciplinary Academic Conference Thailand 2015 organized by International Multidisciplinary Research Foundation, Jomtien, Pattaya on 02-06 November, 2015. (Oral Presentation).