RISK FACTORS ASSOCIATED WITH STILLBIRTH, A CASE STUDY OF THE HOHOE MUNICIPALITY, GHANA.

BY

BREMPONG GUSTAV KUDZO (B.ed Maths, Hons.)

A Project submitted to the Department of Mathematics,

Kwame Nkrumah University of Science and Technology

in partial fulfilment of the requirement for the degree of

MASTER OF SCIENCE

Faculty of Physical Sciences, Department of Mathematics

OCTOBER, 2011

DECLARATION

I hereby declare that this submission is my own work towards the MSc and that, to the best of my knowledge, it contains no materials previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

GUSTAV KUDZO BREMPONG	....	....
20103974		
Student Name & ID	Signature	Date

Certified by:

DR. S.K. AMPONSAH	....	....
Supervisor(s) Name	Signature	Date

Certified by:

MR. K.F. DARKWAH	....	....
Head of Dept. Name	Signature	Date

Certified by:

PROF. I.K. DONTWI	....	....
Dean IDL	Signature	Date

ACKNOWLEGEMENT

I am so grateful to God Almighty for his uncommon favour, guidance, wisdom, knowledge and understanding that he granted me to be able to organize this work.

It would be unfair if I fail to acknowledge my indefatigable supervisor Dr S.K Amponsah of Mathematics Department of KNUST, without whose advice and suggestions this work would not have been completed. In spite of his tight schedules, he was so supportive throughout the organization of this work.

Also worth mentioning is the contribution of Hohoe municipal Health Directorate, Hohoe Municipal Hospital for allowing me to use their institutions for research; and their staff especially midwives at the maternity blocks (Blocks G and H) as well as students on practice from Hohoe midwifery Training School who assisted in the administration of the questionnaires.

Furthermore, my sincere thanks go to Miss Venunye Akpade, a rotational nurse at Hohoe Municipal Hospital who assisted in data collection especially during the visits to the wards and homes.

Not enough could be said to thank my dear wife and children Sylvia, Portia and Patricia for your patience, understanding, and support throughout this long learning effort . You are each a gift to me.

I am also grateful to Mr. Aviah Samuel of Boon Consult who painstakingly typed and helped with the basic editing of the entire thesis.

Lastly, to all relatives and friends, who supported in diverse ways to make this a reality.

.

DEDICATION

This work is dedicated to my children Portia Brempong and Patricia Brempong; my brothers David Brempong and Richard Brempong and the entire family.

Finally, this work is dedicated to all women who never heard their babies cry. Your loss is remembered.

ABSTRACT

Stillbirth is a silent traumatic canker which is a major concern of various individuals, health institutions and the country as a whole. All over the world researchers are fighting tooth and nail to unravel the mystery surrounding the high prevalence of stillbirths. A lot of resources are committed to this area of research because of its alarming rate. This project was carried out in Hohoe Municipality with the aim of finding the risk factors associated with stillbirth through designed questionnaires. Various descriptive analyses were employed on the data obtained. Logistic model was used to identify which independent variables explain the incidence of stillbirth in Hohoe Municipality. The model was based on 250 clients' birth records of mothers who delivered at Hohoe Municipal Hospital from 1st Jan 2011 to 30th June 2011.The predictors of stillbirth were obstetric problem with p-value of 0.002 and odds ratio of 8.540 (95 % C.I. =2.253 – 32.365), place of residence had p-value of 0.044 and odds ratio of 1.178 (95% C.I. = 0.366 – 3.787) alcohol intake recorded p-value of 0.012 and odds ratio 2.064 (95% C.I. = 0. 730 – 5.836), self-medication had p-value of 0.032 and odds ratio of 1.291 (95% C.I. = 0.413 – 4.031). Collinearity diagnostic test conducted showed that the tolerance value of each independent variable was less than 1, signifying that there was no interaction between the variables. Hosmer and Lemshow test of goodness of fit recorded a p-value of 0.542; hence the null hypothesis that the model fits the data well could not be rejected.

TABLE OF CONTENTS

Contents **Page**

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Overview

This chapter contains the background, problem statement, objectives of the study, research questions, scope of the study, justification of the study, delimitations and limitations and organization of the study.

1.1 Background of the Study

Zetterstrom (2007) said most pregnancies are uncomplicated and result in a healthy child and an equally healthy mother who recovers well after delivery. That is what we normally expect. But unfortunately, it is not always the case; the pregnancy can get seriously complicated both for the mother and for the child.

According to Nobel (1979), a high-risk pregnancy is one in which the mother or foetus has a significantly increased chance of death or disability. In order to achieve optimal perinatal outcome, all factors contributing to mortality and morbidity in a particular pregnancy must be identified and acted upon early.

Africa tops preterm births around the world at 11.9%, followed by North America (10.5%) and Asia (9.1%), according to "the Global and regional Toll of Preterm Birth, a report from the March of Dimes Charity based on Statistics published in the Bulletin of the World Health Organization (W.H.O, 2009). Of the 130 million babies born worldwide every year, approximately 4 million are stillborn, more than 98% of these occur in developing countries. Stillbirth accounts for more than half of perinatal mortality in developing countries. In Sub-Saharan Africa, stillbirths account for more than 3% of deliveries each year. While countries in South-East Asia report the highest overall numbers of stillbirth, countries in Africa report the highest incidence rates per 1000 live births. The average stillbirth rate in developing countries has been reported to be 26 per 1000 live births, about five times higher than in developed countries (5 per 1000). One fourth to one third of all stillbirths is estimated to take place during delivery. Stillbirths occurring in the intrapartum period generally have a normal appearance and are often called “fresh” stillbirths. The skin not being intact implies death more than 24 hours before delivery (antepartum), often called “macerated” stillbirths.

Ghana has prioritized child and maternal health care in its development policy framework. Although budgetary allocations to the health sector have increased, causing increase of health facilities, immunization, access to health care services and incidence of Hospital deliveries, impact indicators such as maternal health, life birth and stillbirth figures have not changed so much.

Hohoe Municipal Hospital has shown increasing figures in the rate of stillbirths as depicted by Table 1.1

Table 1.1: Figures on maternal health from Hohoe Municipal Hospital.

Type of Service	2006	2007	2008	2009	2010
Registrants	1562	1781	1998	2319	2245
No. of Attendants	8851	8430	8778	11457	11839
No. of Supervised Deliveries	1364	1405	1706	1882	1935
No. of babies	1413	1449	1776	1928	1985
No. of live babies	1296	1356	1675	1851	1920
No. of Still births(Fresh)	49	31	49	31	27
No. of Still births (Macerated)	67	40	40	44	38
No. of Maternal Death	6	7	8	7	10
No. of Maternal Death Audits	6	7	8	7	10
No. of Abortions	347	359	307	335	230
Post-Natal Attendance	947	1550	1240	1652	1591

Source: 2010 Annual Report of Hohoe Municipal Hospital

1.2 Statement of the Problem

Even though the government of the Republic of Ghana has put in place several measures such as the National Health Insurance Scheme (NHIS), Free Child Delivery Program (FCDP) for pregnant women among other measures in order to meet the Millennium Development Goals (MDG's) on health, the problem of stillbirths continues to prevail in the country at still alarming rates.

As already stated, over 7200 still births are recorded in the world each day. According to a report published by PUDMED, 5% of all deliveries are stillbirths in Ghana. The figures from 2010 report of Hohoe Municipal Health Directorate attest to this fact.

Notwithstanding the efforts being made by the government and other non-governmental agencies to reduce or possibly eradicate stillbirths, the rate at which stillbirths continue to occur in the country has motivated the researcher to undertake much study into this problem and thereby suggest more ways of reducing or eradicating the problem.

1.3 Objective of the Study

The objectives of the study are:

(i) To identify risk factors associated with stillbirths and determine which variables best explain the prevalence of stillbirths in Hohoe Municipality.

(ii) To examine the dependence (if it exists) between individual risk factors identified.

(iii) To determine whether gender is associated with stillbirth

(iv) To use the logistic model to predict the likelihood of a stillbirth using the independent variables identified.

1.4 Research Questions

The research questions of the study are:

(i) What are the most outstanding causes of stillbirths in Hohoe Municipality?

(ii) What is the inter-dependence of independent variables of the occurrence of stillbirths in Hohoe Municipality?

(iii) Is gender associated with stillbirths?

(iv) What statistical model can be used to predict the likelihood of a stillbirth?

1.5 Scope of the Study

The study covers risk factors associated with stillbirths in Hohoe municipality.

Hohoe Municipality is one of the twelve Districts/Municipalities in Volta region and it is located in the central part of the region. It is bounded on the North by Jasikan district South by Ho district East by Republic of Togo and West by Kpando district.

The Municipality covers an area of 1172 sq. km. consisting of 174 communities with a population of 184,743 from the year 2000 National Population census. The population growth rate is 1.9%. The Municipality has been divided into six (6) (Health zones) namely: Akpafu, Alavanyo, Gbledi, Have, Leklebi and Likpe. Hohoe Municipality has a total of 33 health institutions. This includes one mission health centre, two private clinics and one Government hospital. The rest are government health centres and clinics. There are 23 static RCH clinics and 119 outreach clinics. Out of the 27 government health facilities in the district 17 are providing maternity services. See Appendix A for the map of health institutions in the Municipality.

It is a mountainous area to the East on the border with Togo. The highest mountain in Ghana mount Afadjato is found in the Municipality.

The main rivers are Dayi and Koloenu. We also have Wli Waterfall and the Sasa Waterfalls near Alavanyo.

The vegetation is mainly savanna with few parches of swamps and semi-deciduous forest.

The climate of the district is put into three categories. These are the major rainy season from May to July, Minor rainy season from August to October and the Dry or Harmattan season from November to February.

Table 1. 2: Demographic Characteristics Hohoe Municipality.

CATCHMENT POPULATION	**% OF TOTAL POPULATION**	**TOTAL POPULATION**
0-5 years	**20**	**36948.6**
6-14 years	**27**	**49880.61**
15-49 years(Women)	**20**	**36948.6**
15-49 years(Men)	**20**	**36948.6**
50-60 years	**8**	**14779.44**
60+	**5**	**9237.15**
Total	**100%**	**184743**

The people are mainly Ewes interspersed with some Guan speaking people. The main activity of the people in the municipality is farming. Among the produce are maize, rice, cassava yam cocoyam, plantain, banana and vegetables. Few traditional cash crops like cocoa are grown. Besides, Hohoe is the main commercial centre of the Volta region and it is characterised by various shades of commercial activities, making the town a busy one throughout the year.

The municipal capital is linked with second class roads, which are motorable throughout the year. The town roads are in fairly good condition but the link roads are in a deplorable state. Ghana Telecom has commissioned a new telephone digital machine therefore providing a wide range of telephone and broad band internet services both internally and externally to the municipal capital. Other telecom networks such as MTN, TIGO, KASAPA and lately ZAIN have also connected their links with the municipality. A few internet cafes are also available making global communication very easy. Postal services including EMS and Federal Express are available.

There are two local FM stations-Lorlornyo FM and Herritage FM in Hohoe Township for informational purposes.

Hohoe municipality has a number of hotels, Guest Houses, Restaurants and Traditional Catering facilities with a wide variety of menu, both continental and local dishes.

The entire municipality including the hospital is connected to the national grid. The hospital itself has two generating plants, which are used during power outages.

Water supply to the municipality is not much of a problem. The township including the hospital is supplied with pipe-born water whilst the surrounding villages and towns have bore-holes. But some residents use untreated water direct from Dayi river close by.

There are several institutions in the Municipality at various levels on the academic ladder. Below are the details:

Table 1.3: Educational Institutions in Hohoe Municipality

TYPE OF INSTITUTION	STATUS AND NUMBER OF INSTITUTIONS		TOTAL
	Public	Private	
Kindergarten	115	41	156
Primary	125	42	167
Junior High School	80	24	104
Senior High School	10	4	14
Technical Vocational Educational Training	4	2	6
Colleges	2	0	2

General sanitation in the township is not good enough. The disposal of both solid and liquid waste is a major problem. Quite often heaps of un-cleared refuse and stagnant waters are common sights therefore reflecting on the disease patterns of the people.

There is a plywood factory – VOLTA FOREST PRODUCTS which has the capacity to supply products across West Africa and beyond. This company has employed a number of young men and women in the municipality.

Table 1.4: shows Reproductive Health Information of Hohoe Municipality from 2008 to 2010.

		Expected Pregnancies			
Selected indicators	2008	2009	2010		
	Achieved	Achieved	Target	Achieved	%
ANC REG.	4398	4561	5076 70%	4439	60
T./PREGNANCY	631	702	961 13%	690	9.3
3RD TRIM	638	546	591 8%	543	7.3
IPTI	29963	1721	5173 70%	375	41.6
IPT2	1883	1039	5113 70%	1822	24.6
IPT3	827	482	5173 7%	884	11.9
SKILLED DEL.	2327	11981	2956 40%	2856	38.6
SUP.DEL.	3024	1981	3695 50%	253	3.4
LIVE BIRTHS	29931	2518	7390 100%	2852	38.5
MACERATED STILL BITH	46	55	0	49	0.6
FRESH STILL BIRTH	47	38	0	32	0.4

BIRTH WT <2.5KG	328	313	0	360	4.8
BIRTH WT >2.5KG	2695	2267	7390 100%	2573	34.8
MATERNAL DEALTH	8	7	0	10	0.1
MATERNAL DEALTH AUDITED	8	7	100%	10	180%
PNC REG.	2755	2459	3695	2745	37

From 2010 annual report of Hohoe Municipal Hospital, 4439 Pregnant women were registered at the Antenatal Clinics during the period being coverage of 60%. Average number of visits per pregnant woman was 3.9 Out of the total registrants, 690 were teenagers (10-19years)

Out of a total of 3109 deliveries, 253 were supervised and the remaining 2856 were done by skilled personnel. This gives a percentage of 3.4 and 38.6 respectively. TBA deliveries were 181 being 2.4%.

A Total of 2,530 postnatal cases were registered. This forms coverage of 34.2%.

1.6 Justification of the Study

According to the World Health Organization's Opportunities for Africa's Newborns 2006 report, 98% of stillbirths occur in developing countries, especially sub-Saharan Africa, and the stillbirth rate for Ghana is 24 per 1000 deliveries. Even though stillbirths represent a large proportion of perinatal deaths, causes of stillbirths are poorly understood in Ghana.

The eradication of stillbirth in Ghana is the main concern of the general public since procreation ensures continuity of human generation.

The 2008, 2009 and 2010 reports indicate that there has been high prevalence of stillbirth cases in Hohoe Municipality. Our hope for good future in terms of economic stability, good development and better standards of living all depend on how best we can take care of maternal health and subsequently our children and their education. It is my fervent hope that findings of this study would serve as a further directive to policy makers and other stake holders in the eradication of stillbirths and other abnormal birth outcomes.

1.7 Significance of the Study

This study is to identify risk factors associated with stillbirth. The study will help identify the determinants of stillbirth in Hohoe Municipality.

It will also help create the awareness and educate health workers in the Hospitals as well as the communities in the Municipality and alleviate the burden of victims of stillbirth.

Last but not the least, the study will enable the researcher make the necessary recommendations to the Health Services to help reduce or eradicate the high prevalence of stillbirths.

1.8 Limitations of the Study

The original plan of the researcher was to conduct a cohort prospective study. However, data collection ended up in a guided interview of nursing mothers in the Maternity ward after delivery, post-natal clinic and home visits, which was retrospective. This was due to the level of education and painful condition of most of the expectant mothers, the daunting task and resources required in following clients through delivery. Also in some cases, the required information was not seen in the Maternal Health Record Book as a result of few visits to Antenatal Clinic and also may be due to negligence on the part of health workers.

The findings are however limited to deliveries that occurred at Hohoe Municipal Hospital since it absorbs almost all referral cases in the Municipality as well as cases from Krachi, Nkwanta, Kadjebi and Jasikan Districts.

1.9 Organization of the Study

This study was organised into five chapters with each chapter having sub-topics for better comprehension of the study.

Chapter one contains the background, problem statement, objectives of the study, research questions, scope of the study, justification of the study, delimitations and limitations and organization of the study.

Chapter two covers review of related literature. This concerns abstracts on risk factors associated with stillbirths and the use of logistic regression.

Chapter three encompasses methodology applied to the study. It covers data collection procedure and techniques, review of the logistic regression model and other tools used in the analyses of data.

Chapter four covers collection, analyses of data and discussions.

Chapter five brought out findings, conclusions and recommendations of the study.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

This chapter presents related literature on risk factors associated with stillbirths and use of logistic regression.

A stillbirth is defined as the death of a baby in the weeks before birth, or during labour and birth. Cut-off points vary, with the World Health Organization (WHO) defining a stillbirth as the death of a baby at 28 weeks of pregnancy or later, while the UK normally defines stillbirth as a death at 24 weeks or later, and other high-income countries use a cut-off date of 22 weeks. In this report, both definitions are used, with the WHO definition used for international comparisons.

2.1 Risk Factors and Consequences of stillbirths

Yatich et al., (2009) assessed Plasmodium/intestinal helminth infection in pregnancy and other risk factors for stillbirth in Ghana. A cross-sectional study of women presenting for delivery in two hospitals was conducted during November-December 2006. Data collected included socio-demographic information, medical and obstetric histories, and anthropometric measures. Laboratory investigations for the presence of *Plasmodium falciparum* and intestinal helminths, and tests for haemoglobin levels were also performed. From the study, stillbirth rate was relatively high in this population (5%). Most of the stillbirths were fresh and 24% were macerated. When compared to women with no malaria, women with malaria had increased risk of stillbirth (OR = 1.9, 95% CI = 1.2–9.3). Other factors associated with stillbirth were severe anaemia, low serum folate concentration, past induced abortion, and history of stillbirth. The fact that most of the stillbirths were fresh suggests that higher quality intrapartum care could reduce stillbirth rates.

Crape et al., (2007) investigated risk factors associated with stillbirths using personal interviews and medical records abstraction in a hospital-based case control study in Thai Nguyen Province, Vietnam. There were 47 stillbirth cases and 365 controls in this study. Maternal education (≤12 years) (Odds Ratio, OR=3.07; 95% CI=1.19–7.96), from rural communities (OR=2.42; 95% CI=1.16–5.03), primiparous (OR=3.83; 95% CI=1.10–13.40) and lack of prenatal care vitamins (OR=2.56; 95% CI=1.25–5.23) were statistically significant risk factors associated with stillbirth in an age adjusted multivariable model. Their findings suggest that improved maternal health education and care in all communities may reduce the burden of foetal loss in that province.

Say et al., (2006) reported that stillbirth accounts for more than half of perinatal deaths in the developing world and occurs more frequently in developing than in developed countries. Recent reviews and studies of stillbirths have discussed common risk factors and causes for stillbirth in greater detail. However, many stillbirths do not have direct obstetric causes and are referred to as unexplained. They examined the risk factors for stillbirths in Northern Vietnam because not much was known about the epidemiology of adverse pregnancy outcomes in that region.

Study participants were women residents of Thai Nguyen Province who delivered singleton births at Thai Nguyen Central General Hospital in 2004–2005. Stillbirth cases were women who experienced the death of her child in the uterus anytime after 20 weeks gestation. Controls were women who delivered a live-born infant after 37 weeks of pregnancy and identified immediately following a case.

Maternal socio-demographic and lifestyle characteristics, reproductive and medical history, pregnancy morbidities and infant outcomes were obtained through personal interviews and medical records abstraction. They evaluated the following covariates as risk factors for stillbirth using logistic regression (SPSS, Version 13.0): maternal age, education, marital status, occupation, income, ethnicity, type of community, parity, gender, chronic hypertension, chronic diabetes, thyroid disease, chronic kidney disease, asthma, planned pregnancy, prenatal visit, consumption of multivitamins and iron-folic pills during pregnancy, protein urea, toxaemia/preeclampsia during the intrapartum period, pregnancy induced hypertension and anaemia during pregnancy.

A prior and statistically significant risk factors determined in univariate analyses were included in a multifactor model adjusted for maternal age. There were 47 stillbirth cases and 365 live-born controls in this study. Women farmers with not more than 12 years of education, from rural communities, expecting the first child, without adequate prenatal care, or did not consume prenatal vitamins had statistically significant increased odds of experiencing stillbirth. Maternal education with not more than 12 years [Odds ratio (OR=3.07); 95% Confidence interval CI=1.19–7.60, P=0.021)], women from rural communities (OR=2.42; 95% CI=1.16-5.03, P=0.018), primiparous (OR=3.83; 95% CI=1.10-13.40, P=0.035) and no prenatal care vitamins (OR=2.56; 95% CI=1.25 - 5.23, P=0.006) remain statistically significant risk factors associated with stillbirth in an age-adjusted multifactor model. No cases had chronic hypertension, diabetes, or pregnancy induced hypertension. From this study, women of low socioeconomic status in rural communities have higher odds of having stillbirths. The most important risk factors amenable to intervention from a public health perspective are increasing coverage, quality and utilization of prenatal care among women in Thai Nguyen Province, with emphasis on women in rural communities.

A study conducted in Department of Obstetrics and Gynaecology, Leicester Warwick Medical School, Robert Kilpatrick Medical Sciences Building, Leicester Royal Infirmary, Leicester LE2 7LX, UK aimed to calculate the stillbirth rate at each gestation and also determine antenatal factors, which may be associated with unexplained stillbirth in a large

UK teaching hospital. This was a retrospective study of all the stillbirths between January 1995 and October 1998. There were 27 170 births of at least 24 weeks, gestation during the study period. Of these 161 were stillbirths giving a stillbirth rate of 5.9/1000; 149 (92.5%) were antepartum. Eighty-two (50.9%) were unexplained. The "risk" of unexplained stillbirth per 1000 ongoing pregnancies was 0.3 at 24-25 weeks, fell steadily to 0.07 at 30-31 weeks and then rose to a peak of 1.16 at 40-41 weeks. About half (49.2%) of the pregnancies that ended in unexplained stillbirths had a normal antenatal course with no associated factors. Associated factors in the unexplained stillbirth group were identified in 50.8% of cases. The most common was intrauterine growth restriction, identified in 41.5% of cases. The "risk" of stillbirth increases at term. No significant antenatal associated factor, detection of which would aid prevention, could be identified in the majority of cases.

Uma et al., (2011) stated that, stillbirth occurs in about 1 of every 200 pregnancies. Suggestions have been made that it can be prevented through use of increased antenatal surveillance and intervention for risk factors such as advanced maternal age, but there is little supporting evidence for this practice.

The aim of this retrospective study was to identify potential pre-pregnancy risk factors for antepartum stillbirth and to determine whether these risk factors are useful in identifying women at a significantly higher risk for stillbirth at term compared with those in the general population not deemed at increased risk of stillbirth. The study population included a cohort of 174,809 singleton births of least 23 weeks of gestation delivered at 10 institutions in the United States between 2002 and 2008. Of these, 712 were antepartum stillbirths and 174,097 were live births. A subset of 155,629 singleton pregnancies was used to estimate the risk of term antepartum stillbirth. Multivariate analysis assessed the relationship between a potential risk factor and the occurrence of antepartum stillbirth, controlling for other risk factors.

Compared with the live birth controls, factors independently associated with increased risk of antepartum stillbirth were African-American and Hispanic race, advanced maternal age (at least 35–39 years), nulliparity, prepregnancy body mass index of at least 30, pre-existing diabetes, chronic hypertension, smoking, and alcohol use. The risk of antepartum stillbirth increased in women with a history of cesarean delivery or preterm birth after adjustment for other risk factors. For term deliveries among women who were non-Hispanic white, 25 to 29 years of age, normal weight, multiparous, and who had no chronic hypertension or pre-existing diabetes, the baseline risk of term stillbirth was 0.8 per 1000. The adjusted risk of term stillbirth increased when the following individual factors were present: pre-existing diabetes (3.1 per 1000), chronic hypertension (1.7 per 1000), African-American race (1.8 per 1000), maternal age of at least 35 (1.3 per 1000), body mass index at least 30 (1 per 1000), and null parity (0.9 per 1000).These findings show that individual risk factors for antepartum stillbirth are poor predictors of term stillbirth.

Kvinnokliniken et al., in Hulthén-Varli I (2009) reported that the incidence of intrauterine foetal death (from the 28th gestational week) is reported to be 3.6/1000 births in Sweden. Continuous improvements in perinatal care have resulted in a dramatic decrease in early neonatal mortality during the last few decades, but no comparable reduction of antenatal mortality has been observed. A better knowledge of the etiology of stillbirth is imperative to achieve a further decrease in the perinatal mortality rate. Advanced maternal age, smoking during pregnancy, maternal overweight/obesity and low socioeconomic class has been demonstrated to increase the risk for stillbirth. Several studies have been published concerning specific causes of intrauterine foetal death. Abruption of the placenta, some infectious disorders, foetal chromosomal aberration/malformation and maternal disease such as diabetes are some of the conditions that are associated to stillbirth. In order to arrive at a potential cause, a number of investigations have been recommended in the diagnostic evaluation of IUFD.

Edmond et al., (2008) said, in developing countries many stillbirths and neonatal deaths occur at home and cause of death is not recorded by national health information systems. A community-level verbal autopsy tool was used to obtain data on the aetiology of stillbirths and neonatal deaths in rural Ghana. Objectives were to describe the timing and distribution of causes of stillbirths and neonatal deaths according to site of death (health facility or home). Data were collected from 1 January 2003 to 30 June 2004; 20,317 deliveries, 696 stillbirths and 623 neonatal deaths occurred over that time. Most deaths occurred in the antepartum period (28 weeks gestation to the onset of labour) (33.0%). However, the highest risk periods were during labour and delivery (intrapartum period) and the first day of life. Infections were a major cause of death in the antepartum (10.1%) and neonatal (40.3%) periods. The most important cause of intrapartum death was obstetric complications (59.3%). There were significantly fewer neonatal deaths resulting from birth asphyxia in the home than in the health facilities and more deaths from infection. Only 59 (20.7%) mothers of neonates who died at home reported that they sought care from an appropriate health care provider (doctor, nurse or health facility) during their baby's illness. The results from this study highlight the importance of studying community-level data in developing countries and the high risk of intrapartum stillbirths and infectious diseases in the rural African mother and neonate. Community-level interventions are urgently needed, especially interventions that reduce intrapartum deaths and infection rates in the mother and infant.

Edmond et al., (2008) conducted a study which evaluated the diagnostic accuracy of a verbal autopsy (VA) tool in ascertaining the causes of stillbirths and neonatal deaths in rural Ghana and was nested within a community-based maternal vitamin A supplementation trial

(ObaapaVit A trial). All stillbirths and neonatal deaths between 1st January 2003 and 30th June 2004 were prospectively included. Community VAs were carried out within 6 months of death and were classified with a primary cause of death by three experienced paediatricians. The reference standard diagnosis was obtained by the study paediatrician in 4 district hospitals in the study area. There were 20,317 deliveries, 661 stillbirths and 590 neonatal deaths with a VA diagnosis in the study population. A total of 311 stillbirths and 191 neonatal deaths had both a VA and a hospital reference standard diagnosis. The VA performed poorly for stillbirth diagnoses such as congenital abnormalities and maternal haemorrhage. Accuracy was higher for intrapartum obstetric complications and antepartum maternal disease. For neonatal deaths, sensitivity was more than 60% for all major causes; specificity was 76% for birth asphyxia but more than 85% for prematurity and infection. Overall, VA diagnostic accuracy was higher than expected in this rural African setting. Their classification system was based on the expected public health importance of the individual causes of death, differing implications for intervention and the ability to distinguish between the individual causes in low-resource settings. They believe this system was easier to use than traditional approaches and resulted in high precision and accuracy. However, further simplifications are needed to allow use of the World Health Organisation VA in routine child health programmes. The diagnostic accuracy of the VA tool should also be assessed in other regions and in multicentre studies.

Lassey and Obed (2002) determined the trend of concurrent maternal and perinatal mortality at the Korle-Bu Teaching Hospital (KBTH), **Ghana**, and proposed measures for its prevention. A retrospective study, from January 1995 to December 2002, of all concurrent maternal and perinatal deaths in which the woman was 28 weeks' gestation or more (or, if gestational age was not known, the baby weighed 1000 g or more) and died either undelivered or in the perinatal period (within 1 week of delivery) at the KBTH. It was found out that over 8-year study period, there was a total of 93 622 deliveries at the KBTH with 108 concurrent maternal and perinatal mortalities, giving a ratio of 115.4 concurrent maternal and perinatal deaths per 100 000 deliveries. More than 80% of the mothers who died had little or no formal education. Of the 108 mothers, 22 died undelivered. The leading cause of death was a medical condition in pregnancy along with eclampsia/gestational hypertension. Of the 86 delivered mothers, the leading cause of concurrent death was a medical condition in pregnancy. Approximately two-thirds (72/108) of the perinatal deaths were stillbirths. Over the study period, there was a rising trend of the obstetric disaster of losing both mother and baby. There was a rising trend of concurrent maternal and perinatal mortality at the KBTH. It was suggested that a regular antenatal clinic be established with both an internist and obstetrician to jointly see and manage women with medical problems. There was a need for improved and adequate resources to improve outcomes for both mother and baby. A waiver of user fees for maternity services was suggested to improve access for needy and at-risk mothers. Concurrent maternal and perinatal death was seen as the latest negative reproductive health index of the deteriorating socioeconomic situation in developing countries and needed to be tackled decisively.

Say et al., (2006) reported that, stillbirth rate is an important indicator of access to and quality of antenatal and delivery care. Obtaining overall estimates across various regions of the world is not straightforward due to variation in definitions, data collection methods and reporting. They conducted a systematic review of a range of pregnancy-related conditions including stillbirths and performed meta-analysis of the subset of studies reporting stillbirth rates. They examined variation across rates and used meta-regression techniques to explain observed variation. The authors identified 389 articles on stillbirth prevalence among the 2580 included in the systematic review. They included 70 providing 80 data sets from 50 countries in the meta-analysis. Pooled prevalence rates show variation across various subgroup categories. Rates per 100 births are higher in studies conducted in less developed country settings as compared to more developed (1.17 versus 0.50), of inadequate quality as compared to adequate (1.12 versus 0.66), using sub-national sample as compared to national (1.38 versus 0.68), reporting all stillbirths as compared to late stillbirths (0.95 versus 0.63), published in non-English as compared to English (0.91 versus 0.59) and as journal articles as compared to non-journal (1.37 versus 0.67). The results of the meta-regression show the significance of two predictor variables development status of the setting and study quality on stillbirth prevalence. In conclusion, stillbirth prevalence at the community level is typically less than 1% in more developed parts of the world and could exceed 3% in less developed regions. Regular reviews of stillbirth rates in appropriately designed and reported studies are useful in monitoring the adequacy of care. Systematic reviews of prevalence studies are helpful in explaining sources of variation across rates. Exploring these methodological issues will lead to improved standards for assessing the burden of reproductive ill-health.

Mario et al., (2007) identified risk factors for stillbirth in developing countries and measured their impact by calculating the population attributable fraction (PAF) for each risk factor. Systematic review of published studies on risk factors for stillbirth within 3 broadly defined categories: infections, other clinical conditions, and context-dependent conditions such as socioeconomic status, maternal literacy, and receipt of antenatal care. Where statistically significant associations were found between a risk factor and occurrence of stillbirth, the PAF (the proportion of cases occurring in the total population that would be avoided if the exposure was removed) was calculated. A total of 33 studies, conducted in 31 developing countries, were included in the review. The definition of stillbirth varied widely in these studies. Risk factors for stillbirth having a PAF higher than 50% were maternal syphilis, chorioamnionitis, maternal malnutrition, lack of antenatal care, and maternal socioeconomic disadvantage. Maternal syphilis prevention, screening and treatment together with other interventions targeting universal use of antenatal care (that includes screening for syphilis) and improving the socioeconomic conditions including nutritional status of the mother, could effectively contribute towards reducing the unacceptably high burden due to stillbirth in developing countries.

Mavalankar et al., (1991) estimated levels and determinants of perinatal mortality; they conducted a hospital-based surveillance and case-control study, linked with a population survey, in Ahmedabad, India. The perinatal mortality rate was 79.0 per 1000, and was highest for preterm low-birth-weight babies. The case-control study of 451 stillbirths, 160 early neonatal deaths and 1465 controls showed that poor maternal nutritional status, absence of antenatal care, and complications during labour were independently associated with substantially increased risks of perinatal death. Multivariate analyses indicated that socioeconomic factors largely operate through these proximate factors and do not have an independent effect. Estimates of attributable risk derived from the prevalence of exposures in the population survey suggested that improvements in maternal nutrition and antenatal and intrapartum care could result in marked reductions of perinatal mortality.

Kristula et al., (1992) used repeated-measures logistic regression model by evaluating the use of somatic cell counts to identify cows with chronic mastitis as determined by bacteriologic culture. The most important variable to predict whether a cow had chronic mastitis was the average linear score of all of the previous Dairy Herd Improvement Association (DHIA) test dates. The model did not accurately classify cows with chronic mastitis. Results from this study suggested that DHIA centres should be cautious using algorithms of the somatic cell count type to report cows with chronic mastitis to dairy farmers.

Marbury et al., (2011) assessed patterns of alcohol consumption in 12,440 pregnant women by interviewing them at the time of delivery. Only 92 women (0.7 per cent) reported drinking

14 or more drinks per week, with most consuming fewer than 21 drinks per week. In the crude data, alcohol intake of 14 or more drinks per week was associated with a variety of adverse pregnancy outcomes, including low birthweight, gestational age under 37 weeks, stillbirth, and placenta abruptio. After use of logistic regression to control for confounding by demographic characteristics, smoking, parity and obstetric history, only the association of placenta abruptio with alcohol consumption of 14 or more drinks per week remained statistically significant. With the exception of placenta abruptio, alcohol intake of fewer than 14 drinks per week was not associated with and increased risk of any adverse outcome. No association was seen with congenital malformations at any level of alcohol intake.

Stacey et al., (2011) determined whether snoring, sleep position, and other sleep practices in pregnant women are associated with risk of late stillbirth. It was a prospective population based case-control study which was conducted in Auckland, New Zealand. 155 women with a singleton late stillbirth (at least 28 weeks' gestation) without congenital abnormality born between July 2006 and June 2009 and booked to deliver in Auckland participated in the study as cases and 310 women with single ongoing pregnancies and gestation matched to that at which the stillbirth occurred were the controls. Multivariable logistic regression adjusted for known confounding factors. The main outcome measures were maternal snoring, daytime sleepiness (measured with the Epworth sleepiness scale), and sleep position at the time of going to sleep and on waking (left side, right side, back, and other). The results showed that, the prevalence of late stillbirth in this study was 3.09/1000 births. No relation was found between snoring or daytime sleepiness and risk of late stillbirth. However, women who slept on their back or on their right side on the previous night (before stillbirth or interview) were more likely to experience a late stillbirth compared with women who slept on their left side (adjusted odds ratio for back sleeping 2.54 (95% CI 1.04 to 6.18), and for right side sleeping 1.74 (0.98 to 3.01)). The absolute risk of late stillbirth for women who went to sleep on their left was 1.96/1000 and was 3.93/1000 for women who did not go to sleep on their left. Women who got up to go to the toilet once or less on the last night were more likely to experience a late stillbirth compared with women who got up more frequently (adjusted odds ratio 2.28 (1.40 to 3.71)). Women who regularly slept during the day in the previous month were also more likely to experience a late stillbirth than those who did not 2.04 (1.26 to 3.27)). In conclusion, this is the first study to report maternal sleep related practices as risk factors for stillbirth, and these findings require urgent confirmation in further studies.

According to Belsten (2008), stillbirth is one of the most common adverse outcomes of pregnancy. Yet, there is little research that examines common risk factors for stillbirth in terms of the timing of stillbirth. This study examined ten bio-psycho-social variables considered to be risk factors for stillbirth (Maternal Age, Race, Socioeconomic Status, Level of Education, Marital Status, Licit and Illicit Drug Use, Obesity, Diabetes, Hypertension, and Adequacy of Pre-natal Care) and their relationship to the estimated gestational age (EGA) when stillbirth occurs. The nominal, dependent variable, the EGA of stillbirth, was dichotomized into two categories: (1) Early Stillbirth which equalled stillbirth 20 to 28 weeks EGA, and (2) Late Stillbirth, which equalled stillbirth after 28 weeks EGA. Data were collected via a retrospective review of the obstetrical medical records of 231 singleton stillbirths that occurred between January 2000 and December 2005, in two tertiary-care publicly funded hospitals located on Long Island, in New York State. Although the study sample was not a probability sample, rather a population of subjects, the data analysis involved descriptive statistics and two logistic regression models. Due to the sample being treated as a population, the results of the analysis are not statistically inferential in nature, and therefore cannot be freely generalized to other pregnant women in the greater obstetrical population. After obtaining the results of the logistic regression, as the variables Race and Diabetes showed the most variation in relation to the EGA of stillbirth, the association between these risk factors and the dependent variable became the focus of this study. Specifically, the results suggested that in that population black women were at an increased risk (Odds Ratio 2.24) of experiencing an Early Stillbirth instead of a Late Stillbirth compared with women who were not black. The results also suggested that the women in that study diagnosed with diabetes during their pregnancy were at an increased risk (Odds Ratio 2.27) of experiencing a Late Stillbirth instead of an Early Stillbirth compared with women without diabetes during their pregnancy. Although the results of the study represent association, not causation, and cannot be inferred outside of its population, it can be said that the findings suggest that future research might focus on racism and discrimination in relation to the EGA of stillbirth. They also suggest that future research might concentrate on methods of intervention (such as reducing the number of women who are over-weight and obese prior to pregnancy), the goal being a reduction in the prevalence of gestational diabetes as it relates to the EGA of stillbirth. These results may have several policy and program implications. In terms of prevention and early detection of chronic diseases, it would be beneficial to implement universal health insurance as this will ease access to health care for all women throughout their lifetime. It would also be beneficial to develop and implement universal policies and programs to improve working and social conditions (e.g. paid and extended family leave after an infant is born, safe and affordable childcare services, flexible work schedules, and job security) to optimize the long-term health of women, children, and families. Last, but not least, the results suggested that it would be prudent to promote greater investments in women's health in terms of time, money and research.

Reddy et al., (2006) examined the relationship of maternal age with stillbirth risk throughout gestation. A total of 5,458,735 singleton gestations without reported congenital anomalies from the 2001 to 2002 National Centre for Health Statistics perinatal mortality and natality files were analyzed. Hazard rates (risk) of stillbirth (fatal death 20 weeks or longer) were calculated for each week of gestation. The risk of stillbirth at 37 to 41 weeks for women 35 to 39 years old was 1 in 382 ongoing pregnancies and for women 40 years old or older, 1 in 267 ongoing pregnancies. Compared with women younger than 35 years old, the relative risk of stillbirth was 1.32 (95% confidence interval 1.22, 1.43) for women 35 to 39 years old and 1.88 (95% confidence interval 1.64, 2.16) for women 40 years old or older at 37 to 41 weeks. This effect of maternal age persisted despite accounting for medical disease, parity, and race/ethnicity.

Hossain and Khan (2009) conducted a study to determine the obstetric causes for stillbirth in low socio-economic settings. A case-control retrospective study through data analysis was conducted at a tertiary university hospital, from January to June 2008. All pregnant women diagnosed with stillbirth after 28 weeks of gestation were included in the study. They were compared with women who had live birth during the study period. Both groups were identified from the admission, and labour room registers. The risk factors studied were maternal age, parity, gestational age, hypertensive disorders of pregnancy, antepartum haemorrhage, obstructed labour and Prematurity. Stillbirth was defined as foetal death after 28 weeks of gestation. Of the 1011 deliveries in the selected period, there were 100 still births (98/1000 deliveries). Both nulliparity and grand multiparity were significantly associated with stillbirths ($p < 0.003$ and $p < 0.009$ respectively). From the binary logistic regression analysis, obstetric factors which were significantly associated with stillbirth were obstructed labour (OR 16.2, CI 5.5-47), hypertensive disorders (OR 9.6 CI 4-23), abruptio placentae (OR 136, CI 52-356), placenta previa (OR 71, CI 21-230), and preterm labour (OR 15 CI 4-54). Gender was not found significantly associated with stillbirth ($p < 0.432$). The study concludes that**,** majority of stillbirths were due to risk factors which can be identified in the antenatal period.

Kesmodel et al., (2001) evaluated the association between alcohol intake during pregnancy and risk of stillbirth and infant death in a cohort of pregnant women receiving routine antenatal care at Aarhus University Hospital (Aarhus, Denmark) between 1989 and 1996. Prospective information on alcohol intake, other lifestyle factors, maternal characteristics, and obstetric risk factors was obtained from self-administered questionnaires and hospital files, and 24,768 singleton pregnancies were included in the analyses (116 stillbirths, 119 infant deaths). The risk ratio for stillbirth among women who consumed at least 5 drinks/week during pregnancy was 2.96 (95% confidence interval: 1.37, 6.41) as compared with women who consumed less than 1 drink/week. Adjustment for smoking habits, caffeine intake, age, pre-pregnancy body mass index, marital status, occupational status, education, parity, and sex of the child did not change the conclusions, nor did restriction of the highest intake group to women who consumed 5–14 drinks/week (risk ratio = 3.13, 95% confidence interval: 1.45, 6.77). The rate of stillbirth due to fetoplacental dysfunction increased across alcohol categories, from 1.37 per 1,000 births for women consuming less than 1 drink/week to 8.83 per 1,000 births for women consuming at least 5 drinks/week. The increased risk could not be attributed to the effect of alcohol on the risk of low birth weight, preterm delivery, or malformations. There was little if any association between alcohol intake and infant death.

Johnson et al., (2006) sought to estimate the association between prenatal smoking and stillbirth in a longitudinal cohort using two study designs: a case–control study and a bidirectional case-crossover study. The analysis was conducted using the Missouri maternally linked cohort dataset from 1978 through 1997. In the case–control study, each mother contributed only one birth to the analysis. For the bidirectional crossover design, analysis was restricted to women who gave birth to at least one stillbirth, and the controls comprised all live births before and after the stillbirth. The independent association between prenatal smoking and stillbirth was computed using non-conditional (case–control design) and conditional (bidirectional case–crossover design) logistic regression. Prenatal smoking decreased from 29.7% in 1978 to 21.2% by 1997 ($p<.001$). The absolute risk of stillbirth was greater among smokers (7.7/1000) than non-smokers (5.3/1000), $p<.001$. In the case–control design, the risk of stillbirth was 34% greater among smokers than non-smokers (*OR*=1.34, 95% *CI* 1.26–1.43). For each 10-unit increase in the number of cigarettes consumed per day prenatally, the likelihood of stillbirth rose by about 14% ($p<.001$). In the bidirectional case–crossover design, the association between stillbirth and smoking during pregnancy was confirmed, although the magnitude of the relationship was smaller (*OR*=1.20, 95% CI 1.03–1.39). In conclusion, they found prenatal smoking to be a risk factor for stillbirth even after minimizing the influence of known and unknown sources of confounding as well as changes in temporal trend in prenatal smoking.

Smith et al., (2007) determined whether maternal serum levels of alpha-fetoprotein (alpha-FP) and human chorionic gonadotrophin (hCG) at 15-21 weeks provided clinically useful prediction of stillbirth in first pregnancies. The design was Retrospective study of record linkage of a regional serum screening laboratory to national registries of pregnancy outcome and perinatal death. It was in West of Scotland, 1992-2001. A total of 84769 eligible primigravid women delivering an infant at or beyond 24 weeks of gestation were considered. The risk of stillbirth between 24 and 43 weeks was assessed using the Cox proportional hazards model. Logistic regression models within gestational windows were then used to estimate predicted probability. Screening performance was assessed as area under the receiver operating characteristic (ROC) curve. The main outcome measure was Antepartum stillbirth unrelated to congenital abnormality. The odds ratio (95% CI) for stillbirth at 24-28 weeks for women in the top 1% were 11.97 (5.34-26.83) for alpha-FP and 5.80 (2.19-15.40) for hCG. The corresponding odds ratios for stillbirth at or after 37 weeks were 2.44 (0.74-

8.10) and 0.79 (0.11-5.86), respectively. Adding biochemical to maternal data increased the area under the ROC curve from 0.66 to 0.75 for stillbirth between 24 and 28 weeks but only increased it from 0.64 to 0.65 for stillbirth at term and post-term. Women in the top 5% of predicted risk had a positive likelihood ratio of 7.8 at 24-28 weeks, 3.7 at 29-32 weeks, 5.1 at 33-36 weeks and 3.4 at 37-43 weeks, and the corresponding positive predictive values were 0.97, 0.33, 0.47 and 0.63%, respectively. In conclusion, maternal serum levels of alpha-FP and hCG were statistically associated with stillbirth risk. However, the predictive ability was generally poor except for losses at extreme preterm gestations, where prevention may be difficult and interventions have the potential to cause significant harm.

Haberman and Sinharay (2010) said most automated essay scoring programs use a linear regression model to predict an essay score from several essay features. This article applied a cumulative logit model instead of the linear regression model to automated essay scoring. Comparison of the performances of the linear regression model and the cumulative logit model was performed on a large variety of data sets. It appears that the cumulative logit model performed somewhat better than did the linear regression model.

Högberg and Cnattingius (2007) said, maternal smoking has previously been associated with risk of stillbirth. If women who quit smoking reduce their risk of stillbirth, the hypothesis of a causal association would be supported. The study was a prospective cohort study. This was conducted Nationwide in Sweden using all primiparous women who delivered their first and second consecutive single births between 1983 and 2001, giving a total number of 526,691 women. A population-based Swedish study with data from the Medical Birth Registry, the Immigration Registry and the Education Registry. Logistic regression analyses were used to estimate odds ratios, using 95% confidence intervals. The main outcome measure was stillbirth in the second pregnancy. Compared with non-smokers in both pregnancies, women who smoked during the first pregnancy but not during the second do not have an increased risk of stillbirth (OR 1.02; 95% CI 0.79-1.30), while corresponding risk among women who smoked during both pregnancies was 1.35 (95% CI 1.15-1.58). The result supports that maternal smoking during pregnancy is causally associated with stillbirth risk. Smoking is a preventable cause of stillbirth, and smoking interventions is an important issue in antenatal care.

KEATING, and CHERRY (2004) stated that logistic regression is an important tool for wildlife habitat-selection studies, but the method frequently has been misapplied due to inadequate understanding of the logistic model, its interpretation, and the influence of sampling design. To promote better use of this method, the study reviewed its application and interpretation under 3 sampling designs: random, case–control, and use–availability. Logistic regression is appropriate for habitat use–nonuse studies employing random sampling and can be used to directly model the conditional probability of use in such cases. Logistic regression also is appropriate for studies employing case–control sampling designs, but careful attention is required to interpret results correctly. Unless bias can be estimated or probability of use is small for all habitats, results of case–control studies should be interpreted as odds ratios, rather than probability of use or relative probability of use. When data are gathered under a use–availability design, logistic regression can be used to estimate approximate odds ratios if probability of use is small, at least on average. More generally, however, logistic regression is inappropriate for modelling habitat selection in use–availability studies. In particular, using logistic regression to fit the exponential model of Manly et al., (2002:100) does not guarantee maximum-likelihood estimates, valid probabilities, or valid likelihoods. They show that the resource selection function (RSF) commonly used for the exponential model is proportional to a logistic discriminant function. Thus, it may be used to rank habitats with respect to probability of use and to identify important habitat characteristics or their surrogates, but it is not guaranteed to be proportional to probability of use. Other problems associated with the exponential model also are discussed. The authors described an alternative model based on Lancaster and Imbens (1996) that offers a method for estimating conditional probability of use in use–availability studies. Although promising, this model fails to converge to a unique solution in some important situations. Further work is needed to obtain a robust method that is broadly applicable to use–availability studies.

2.2 Application of Logistic Regression model

Flom (2002) stated that, logistic regression may be useful when we are trying to model a categorical dependent variable (DV) as a function of one or more independent variables. Their paper reviewed the case when the DV has more than two levels, either ordered or not, gave and explained SASR code for these methods, and illustrated them with examples.

In logistic regression, the goal is the same as in ordinary least squares (OLS) regression: the authors modelled a dependent variable (DV) in terms of one or more independent variables (IVs). However, OLS regression is for continuous (or nearly continuous) DVs; logistic regression is for DVs that are categorical. The DV may have two categories (e.g., alive/dead; male/female; Republican/ Democrat) or more than two categories. If it has more than two categories they may be ordered (e.g. none/some/a lot) or unordered (e.g. married/single/ divorced/widowed/other). Their paper dealt with modelling multiple category DVs (ordered or not) with SAS PROC LOGISTIC.

WHY LOGISTIC REGRESSION

One might try to use OLS regression with categorical DVs. There are several reasons why this is a bad idea:

(i) The residuals cannot be normally distributed (as the OLS model assumes), since they can only take on one of several values for each combination of level of the IVs

(ii) The OLS model makes nonsensical predictions, since the DV is not continuous - e.g., it may predict that someone does something more than 'all the time'.

(iii). For nominal DVs, the coding is completely arbitrary, and for ordinal DVs it is (at least supposedly) arbitrary up to a monotonic transformation. Yet recoding the DV will give very different results.

INTRODUCTION TO LOGISTIC REGRESSION

Logistic regression deals with these issues by transforming the DV. Rather than using the categorical responses, it uses the log of the odds ratio of being in a particular category for each combination of values of the IVs. The odds is the same as in gambling, e.g., 3-1 indicates that the event is three times more likely to occur than not. We take the ratio of the odds in order to allow us to consider the effect of the IVs. We then take the log of the ratio so that the final number goes from $-\infty$ to $+\infty$, so that 0 indicates no effect, and so that the result is symmetric around 0, rather than 1. For more details on logistic regression, see Hosmer and Lemeshow (2000), Agresti (2002), or Long (1997).

Methods such as forward, backward, and stepwise selection are available, but, in logistic as in other regression methods are not to be recommended. They give incorrect estimates of the standard errors and p-values, can delete variables that are critical to include, and, perhaps most important, allow the researcher not to think (Harrell, 2001). It is much better to compare models based on their results, reasonableness, and fit (as measured, e.g. by the Akaike Information Criterion (AIC)—note that a lower AIC indicates better fit). A good text on this is Burnham and Anderson (2002).

Hale (2001) described an application of logistic regression to mapping of gold potential in the Baguio district of the Philippines. Categorical map data such as litho-logic units and proximity classes of curvi-linear features, based on spatial association analyses, are quantified systematically and used as independent variables in logistic regression to predict the probability for presence or absence of gold mineralization. Regression experiments to compare between using all independent variables that are associated spatially with the response variable and using only statistically significant independent variables are performed. The results of the regression experiments are similar; however, the use of all independent variables produces slightly optimistic results but better prediction rates for the known gold deposits in the test district. At least 68% of the 'model' large-scale gold deposits and at least 76% of the 'validation' small-scale gold deposits were predicted correctly. The predicted geologically favourable zones are also similar to delineated geochemically anomalous zones. The technique presented using logistic regression as a data integration tool is effective for geologically constrained technique of mapping mineral potential.

ZHU and LI (2010) in their study applied logistic regression to credit risk analysis. With the appearance of listed companies' credit issues and frequent credit crisis, investors are increasingly concerned about credit risk analysis for listed companies. In view of the current development methods of credit risk analysis and the importance of identifying corporate financial risk, this paper designed an effective indicator system and established the credit evaluation models of China's listed companies by taking advantage of their 2009 financial data. Combined with the reality of China's listed companies, they used the established models to discriminate and analyze. The result of empirical research on the credit risk analysis for listed companies is that Logistic regression model is superior to discriminant analysis model.

Hung and Wang (2011) said logistic regression has been widely applied in the field of biostatistics for a long time. It aims to model the conditional success probability of an event of interest as the logit function of a linear combination of covariates, for the sake of further interpretation of covariates and prediction of new observation. In some applications, however, covariates of interest have a natural structure, such as being a matrix, at the time of being collected. The rows and columns of the covariate matrix would have different meanings, and they must contain useful information regarding the response. If we simply stack X as a vector and fit the conventional logistic regression model, we may discard relevant information and may also suffer the problem of inefficiency in estimating parameters. Motivated from this reason, they propose in this paper the matrix variate logistic (MV-logistic) regression model. The most important feature of their model is that it retains the inherent structure of the covariate matrix. Another advantage is the parsimony of parameters needed. These features lead to a good performance of MV-logistic regression in many situations. Simulation studies and a data example, the EEG data, demonstrate the usefulness of the proposed method.

Bayaga (2010) explored the usage of multinomial logistic regression (MLR) in risk analysis. In this regard, performing MLR on risk analysis data corrected for the non-linear nature of binary response and did address the violation of equal variance and normality assumptions. Additionally, use of maximum likelihood (-2log) estimation provided a means of working with binary response data. The relationship of independent and dependent variables was also addressed. The data used included a cohort of hundred risk analyst of a historically black South African University. In this analysis, the findings revealed that the probability of the model chi-square (17.142) was 0.005, less than the level of significance of 0.05 (i.e. $p<0.05$). Suggesting that there was a statistically significant relationship between the independent variable-risk planning (Rp) and the dependent variable-control mechanism (control mecs) ($p<0.05$). Also, there was a statistically significant relationship between key risks assigned (KSA) and time spent on risk mitigation. For each unit increase in confidence in control mecs, the odds of being in the group of survey respondents who thought institution spend too little time on Rp decreased by 74.7%. Moreover, the findings revealed that survey respondents who had less confidence in control mecs were less likely to be in the group of survey respondents who thought institution spent about the right amount of time on risk planning.

Kayri* and Çokluk (2010) studied nation of Artificial Neural Networks' classification and parameter estimation with Multinomial Logistic Regression Analysis. One of the modeling types suggested in case of having the dependent variables in categorized/classified structure and the independent variables in different structures such as nominal, ordinal, and intervals etc. in a research pattern is "Multinomial Logistic Regression (MLR)" method. MLR and Artificial Neural Networks (ANN) based MLR Analyses' findings were studied comparatively in the model, where the dependent variable performed categorical structure and the independent variables performed mixed (continuous-discrete) structure. For the research, real data that were gathered in the context of the study entitled "Studying Primary School Students' Views on their Communications with the Teachers and the Expected Situation" were used by the "Students' Expectations from their Teachers in Teacher-Student Communication Process Scale" developed by Doğan (2009). Within the context of this study, the total score obtained from the scale was assigned as the dependent variable, and variables such as the school type (public-private), gender, grade, mother's profession, father's profession, mother's educational status, father's educational status, number of brothers/sisters, monthly income, and internet usage time were assigned as independent variables. ANN has classified the dependent variable in high correctness level and showed the model's fit in a higher level than MLR. Moreover, ANN has obtained parameter coefficients unlike MLR. It was considered that the model studied was estimated more consistently and correctly with ANN. Scientific researches not only aim at describing a current situation, but also constantly bringing up cause and effect relations between actions. From this point of view, it was stated that the scientific researches try to present the cause and effect relations in a model and that they realize these models in different methods according to the structures of the data. Result(s) is taken as dependent variable, and the set of causes affecting the result as independent variables in cause and effect based researches. Modelling also vary according to the data structures of the dependent and independent variables.

Linear regression models (univariate linear, multiple, and multivariate linear) are commonly used models in social sciences; these models require the dependent and independent variables to be constant variables. Moreover, they also require meeting assumptions such as the distribution normality of dependent and independent variable sets, and error distribution normality of observation values. On the contrary, the results obtained in social sciences researches are generally categorical and so it is not possible to study them within linear regression models' context. Accordingly, one of the regression models that can be used in such cases is logistic regression. Logistic regression is an analysis type appropriate for the situations where the dependent variable is not continuous or quantitative, in other words, it is appropriate for categorical or nominal situations (Long, 1997; Mertler & Vannatta, 2005). As a result of this, logistic regression does not require meeting basic assumptions in linear regression models.

Derr (2009) said, exact logistic regression has become an important analytical technique, especially in the pharmaceutical industry, since the usual asymptotic methods for analyzing small, skewed, or sparse data sets are unreliable. Inference based on enumerating the exact distributions of sufficient statistics for parameters of interest in a logistic regression model, conditional on the remaining parameters, is computationally infeasible for many problems. Efficient algorithms for generating the required conditional distributions were introduced in Hirji, Mehta, and Patel (1987) and Mehta, Patel, and Senchaudhuri (1992, 2000), thus making these methods computationally available. The paper discussed the theory and methods for exact logistic regression and illustrates their application with the LOGISTIC procedure in SAS/STAT® 9.2 software.

Many clinical trials deal with the comparison of populations of subjects with categorical responses. Historically, statistical inference for such studies involved large-sample approximations, and fitting logistic regression models to such data was performed through the unconditional likelihood function. However, asymptotic methods might be inadequate when sample sizes are small or the data are sparse, skewed, or heavily tied. Exact conditional inference remains valid in such situations.

The LOGISTIC, GENMOD, GLIMMIX, PROBIT, and CATMOD procedures perform unconditional likelihood inference for logit models, and the LOGISTIC and PHREG procedures can perform asymptotic conditional likelihood inference for logit models. SAS users have requested the ability to perform exact tests for logistic regression modeling. Many exact statistical tests have already been added to the FREQ and NPAR1WAY procedures, and as of SAS 8.1, SAS/STAT software includes exact logistic regression for binary (dichotomous) response variables in the LOGISTIC procedure. Exact methods for generalized logit (GLOGIT) models have been available in the LOGISTIC procedure since SAS 9.

The "METHODOLOGY" section in this paper presented the logistic regression model and the different likelihoods, and then explained how the exact analysis algorithm implemented in PROC LOGISTIC works.

2.3 Strength and Weakness of Logistic Regression Model

Lee (1985) said multiple logistic regression is an accepted statistical method for assessing association between an anticedant characteristic (risk factor) and a quantal outcome (probability of disease occurrence), statistically adjusting for potential confounding effects of other covariates. Yet the method has potential drawbacks which are not generally recognized. This article considers one important drawback of logistic regression. Specifically the so-called main effect logistic model assumes that the probability of developing disease is linearly and additively related to the risk factors on the logistic scale. This assumption stipulates that for each risk factor, the odds ratio is constant over all reference exposure levels, and that the odds ratio exposed to two or more factors is equal to the product of individual risk factor odds ratios. If the observed odds ratios in the data follow this pattern, the model-predicted odds ratios will be accurate, and the meaning of the odds ratio for each risk factor will be straightforward. But if the observed odds ratios deviate from the model assumption, the model will not fit the data accurately, and the model-predicted odds ratios will not reflect those in the data. Although satisfactory fit can always be achieved by adding to the model polynomial and product terms derived from the original risk factors, the odds ratios estimated by such an interaction logistic model are difficult to interpret, viz., the odds ratio for each risk factor depends not only on the reference exposure levels of that factor, but also on the exposure level in other factors. Empirical evidence suggests that the actual relationship between risk factors and disease is likely to be nonlinear and non additive on the logistic scale. Consequently, in most instances, an interaction logistic model is required to analyse epidemiological data satisfactorily. The expressed purpose of this article was to make explicit the potential problems encountered in logistic regression analysis. It is hoped that an awareness of these subtleties will encourage a more judicious use of this method in the analysis of epidemiological data.

CHAPTER THREE

METHODOLOGY

3.0 Introduction

This chapter looks at the methods used in the study.

3.1 Population and Sample Selection

Hohoe Municipal Hospital had 1073 deliveries from 1st January, 2011 to 30th June, 2011. By convenience and time constraint, a sample size of 250 clients was chosen. This represents about 23% of the total deliveries during the period. To ensure that each client had equal chance of being selected, the researcher paid random visits to the maternity wards after delivery and those at postnatal clinic. Some of the clients who had stillbirth but were not met at the hospital were traced to their homes.

3.2 Data Collection Procedure

Secondary data was not available in the form that would suit a logistic model. The researcher collected primary data using personal interview with structured questionnaire designed to meet the objectives of the study. See Appendix C for the questionnaire.

The initial intention was to follow the expectant mothers till the outcome of pregnancy is seen, but this became very daunting and impossible in certain cases. Instead, the researcher met clients in the maternity wards and those who reported for post-natal services. Some of the clients who had stillbirth were visited at home.

3.3 Method of Analyses

3.3.1 Characteristics of the Respondents

The preliminary steps of the editing, coding and tabulating of the data were employed. The administered questionnaire was edited to impose some minimum quality standard on the raw data. The raw data was categorised according to the various variables of respondents.

Among the variables are the mother's age, level of education, source of water for drinking, type of fuel wood the mother used for cooking, type of residence, number of antenatal visit, maternal food intake, maternal health status, household size, maternal height, maternal weight gain during pregnancy, employment status, occupation, income level, history of previous stillbirth, birth spacing, sex of new born, weight of baby and anaemia in terms Hb.

Also, factors on maternal behaviours such as smoking (including passive smoking), substance abuse (drugs), alcohol intake, use of colanuts or white clay and number of abortions (both spontaneous and induced).

Logistic regression model was used to determine the risk factors associated with stillbirths by estimating the odds ratios (OR) and their 95% confidence intervals (CI). After identifying all the causal variables that showed statistical significance in the univariable analysis, a multiple analysis was conducted to control for confounders.

3.4 Some Basic Concepts

3.4.1 Objectives of the Logistic Regression Model

Logistic regression is a mathematical modelling approach that can be used to describe the relationship of several *X'*s to a dichotomous dependent variable, such as *D*.

Other modelling approaches are possible also, but logistic regression is by far the most popular modelling procedure used to analyze epidemiologic data when the illness measure is dichotomous.

3.4.2 Why Is Logistic Regression Popular?

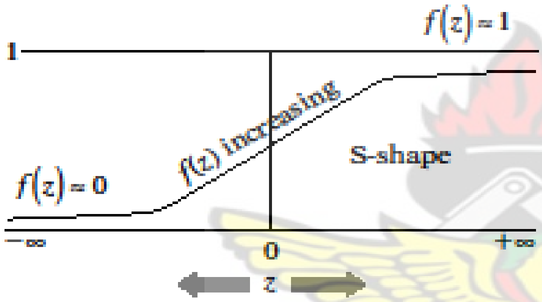
To explain the popularity of logistic regression, we show here the logistic function, which describes the mathematical form on which the logistic model is based. This function, called $f(z)$, is given by 1 over 1 plus e to the minus z. We have plotted the values of this function as z varies from $-\infty$ to $+\infty$.

Independent variables: $X_1, X_2, \ldots, X_k$

X's may be E's, C's, or combinations

3.4.3 The Logistic Function

EXAMPLE



$$f(z) = \frac{1}{1 + e^{-z}}$$

Fig 3.1 Graph showing the non-linearity of the logistic regression function

The fact that the logistic function $f(z)$ ranges between 0 and 1 is the primary reason the logistic model is so popular. The model is designed to describe a probability, which is always some number between 0 and 1. In epidemiologic terms, such a probability gives the risk of an individual getting a disease.

3.4.4 The Logistic Model

The logistic model, therefore, is set up to ensure that whatever estimate of risk we get, it will always be some number between 0 and 1. Thus, for the logistic model, we can never get a risk estimate either above 1 or below 0. This is not always true for other possible models, which is why the logistic model is often the first choice when a probability is to be estimated. Another reason why the logistic model is popular derives from the shape of the logistic function. As shown in the graph, if we start at $z = -\infty$ and move to the right, then as z

increases, the value of *f*(*z*) hovers close to zero for a while, then starts to increase dramatically toward 1, and finally levels off around 1 as *z* increases toward $+\infty$.The result is an elongated, S shaped picture.

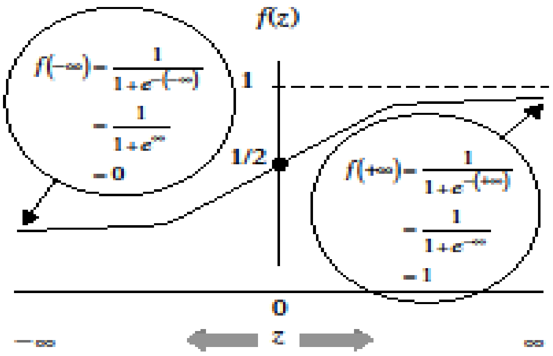


Fig 3.2 A plot of f(z) as z takes values from $-\infty$ *to* $+\infty$

Now, let's go from the logistic function to the model, which is our primary focus. To obtain the logistic model from the logistic function, we write *z* as the linear sum $\propto$ plus β_1 times X_1 plus $\beta_2$times X_2, and so on to β_k times X_k, where the *X*'s are independent variables of interest and $\propto$ and the β_i are constant terms representing unknown parameters. In essence, then, *z* is an index that combines the *X*'s. We now substitute the linear sum expression for *z* in the right-hand side of the formula for *f*(*z*) to get the expression *f*(*z*) equals 1 over 1 plus *e* to minus the quantity $\propto$ plus the sum of βiX*i* for *i* ranging from 1 to *k*. Actually, to view this expression as a mathematical model, we must place it in an epidemiologic context.

$$\mathbf{z = \alpha + \beta_1 x_1 + \ \beta_2 x_2 + \cdots, +\beta_k x_k}$$

$$\mathrm{f(z)} = \frac{1}{1 + \mathrm{e}^{-\mathrm{z}}}$$

$$= \mathrm{f(z)} = \frac{1}{1 + \mathrm{e}^{-(\alpha + \sum \beta_i x_i)}}$$

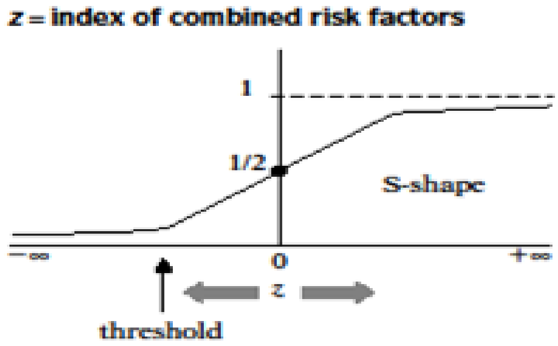


Fig 3.3 Graph showing the threshold concept of disease conditions

3.4.4.1 Epidemiologic Framework

The logistic model considers the following general epidemiologic study framework**:** We have observed independent variables X_1, X_2, and so on up to X_k on a group of subjects, for whom we have also determined disease status, as either 1 if "with disease" or 0 if "without disease."

$X_1, X_2, \ldots, X_k$ measured at T_0

Time: $T_O \rightarrow T_1$

$$X_1, X_2, \ldots . X_k \rightarrow D_{(0,1)}$$

$$\mathrm{P}(D=1|\ X_1, X_2, \ldots, X_k)$$

We wish to use this information to describe the probability that hte disease will develop during a defined study period, say say T_0 to T_1, in a disease-free individual with independent variable values X_1, X_2, up to X_k which are measured at T_0.The probability being modelled can be denoted by the conditional probability statement $\mathrm{P}(D=1|X_1, X_2,\ldots, X_k)$.

The model is defined as logistic if the expression for the probability of developing the disease, given the *X'*s, is 1 over 1 plus *e* to minus the quantity α plus the sum from *i* equals 1 to *k* of β_i times *Xi*.

The terms α and β_i in this model represent unknown parameters that we need to estimate based on data obtained on the *X'*s and on *D* (disease outcome) for a group of subjects. Thus, if we knew the parameters and α the βi and we had determined the values of X_1 through X_k for a particular disease-free individual, we could use this formula to plug in these values and obtain the probability that this individual would develop the disease over some defined follow-up time interval. For notational convenience, we will denote the probability statement P(D=1| X_1, X_2, . . . , X_k) as simply P(X) where the bold X is a shortcut notation for the collection of variables X_1 through X_k. Thus, the logistic model may be written as P(**X**) equals 1 over 1 plus e to minus the quantity α plus the sum βiXi.

Logistic model is given as

P(D=1=X_1, X_2, . . . , X_k) = $\frac{1}{1+e^{-(\alpha+\sum\beta iXi)}}$

NOTATION

P (D=1|X_1, X_2, . . . , X_k) = P(X)

Thus the Model formula is P(X)= $\frac{1}{1+e^{-(\alpha+\sum\beta iXi)}}$

In practice non-linear relation between P(x) and x are often monotonic with P(x) increasing continuously or P(x) decreasing continuously as x increases. This is shown by the S-shaped curve in fig.........As $x \to \infty, P(x) \downarrow 0 \; when \; \beta < 0, P(x) \uparrow 1 \; when \; \beta > 0$

3.4.4.2 ODDS RATIO

To begin the description of the odds ratio in logistic regression, we present an alternative way to write the logistic model, called the logit form of the model. To get the logit from the logistic model, we make a transformation of the model.

3.4.4.3 The Logit Transfomatiom

The logit transformation, denoted as logit P(X), is given by the natural log (i.e., to the base e) of the quantity P(X) divided by one minus P(X), where P(X) denotes the logistic model as previously defined.

This transformation allows us to compute a number, called logit P(X), for an individual with independent variables given by X. We do so by:

(i) computing P(X) and

(ii) 1 minus P(X) separately, then

(iii) dividing one by the other, and finally

(iv) taking the natural log of the ratio.

Logit P(X) = $\ln_e\left[\frac{P(X)}{1-P(X)}\right]$, where P(X) = $\frac{1}{1+e^{-(\alpha+\sum\beta iXi)}}$

We get a general formula when we plug the logistic model into the logistic function. This formula would have a relationship with the odds ratio as shown below.

$$P(\mathbf{X}) = \frac{1}{1+e^{-(\alpha+\sum\beta iXi)}}$$

$$1- P(\mathbf{X}) = 1- \frac{1}{1+e^{-(\alpha+\sum\beta iXi)}} = \frac{e^{-(\alpha+\sum\beta iXi)}}{1+e^{-(\alpha+\sum\beta iXi)}}$$

$$\frac{P(X)}{1-P(X)} = \frac{\frac{1}{1+e^{-(\alpha+\sum\beta iXi)}}}{\frac{e^{-(\alpha+\sum\beta iXi)}}{1+e^{-(\alpha+\sum\beta iXi)}}} = e^{(\alpha+\sum\beta iXi)}$$

$$\boldsymbol{ln}\left[\frac{\boldsymbol{P(X)}}{\boldsymbol{1-P(X)}}\right] = \ln[e^{(\alpha+\sum\beta iXi)}] = \alpha + \sum\beta iXi$$

Hence in logit form,
Logit P(X) = $\alpha + \sum\beta iXi$

Whereas f(z) must fall in the (0,1) range, the logit can be any real number. The real numbers are also the range for the linear predictors $(\alpha + \sum\beta iXi)$

3.5 Applicability of logistic modelling in statistical research

Binary logistic regression is used to analyze relationships between a dichotomous dependent variable and metric or dichotomous independent variables.

Logistic regression combines the independent variables to estimate the probability that a particular event will occur, that is, a subject will be a member of one of the groups defined by the dichotomous dependent variable. The predicted event for a particular analysis is referred to as the modelled event. More generally, the independent variables can be denoted as $X_{1,}X_2$ and so on up to X_k where k is the number of variables being considered. The X's, can represent any collection of exposure variables or risk factors, even combinations of such variables of interest.

3.6 The multivariable problem

Whenever we wish to relate a set of X's $(X_1, \; X_2, \ldots, X_k)$ to a dependent variable, say, Y, we are considering a multivariable problem. In the analysis of such a problem, some kind of mathematical model is typically used to deal with the complex interrelationships among the many variables. Logistic regression is a mathematical modelling approach that can be used to describe the relationship of several X's to dichotomous dependent variable, such a Y, by using the multiple logistic regression model.

3.7 What logistic regression predicts.

The variate or value produced by logistic regression is a probability value between 0 and 1. If the probability for group membership in the modelled category is above 0.50, the subject is predicted to be a member of the modeled group. If the probability is below 0.5, the subject is predicted to be a member of the other group.

For any group case, logistic regression computes the probability that a case with a particular set of values for the independent variable is a member of the modelled category.

3.8 Measurement requirements

Binary logistic regression analysis requires that the dependant variable be dichotomous.

Binary logistic regression analysis also requires the independent variables be metric or dichotomous.

If an independent variable is nominal level and not dichotomous, the logistic regression procedure in SPSS has an option to dummy code the variable.

3.9 Assumptions

Logistic regression does not make any assumptions of normality, linearity, and homogeneity of variance for the independent variables.

Because it does not impose these requirements, it is preferred to other types of analysis such as the discriminant analysis when the data does not satisfy these assumptions.

3.10 Sample size considerations

Sample size calculation for logistic regression is a complex problem, but on the work of Pedizzi et al (1996), the following guideline for a minimum number of cases to include in the study is suggested.

Let p be the smallest of the proportions of negative or positive cases in the population and k the number of covariates (the number of independent variables), then the minimum number of cases to include is: $N = \frac{10k}{p}$

Alternatively, using a guideline provided by Hosmer and Lemshow, 2000), the minimum number of cases per independent variable is 10, with a preferred ratio of 20 to 1.

3.11 Methods for including variables

There are three methods available for including variables in the regression equation: the simultaneous method in which all independent variables are included at the same time, hierarchical method in which control variables are entered in the analysis before predictors whose effects we are primarily concerned with the stepwise method in which variables are selected in the order in which they maximize the statistically significant contribution to the model.

For all methods, the contribution to the model is measured by the model likelihood ratio statistic (LR), a statistical measure of the fit between the dependent and independent variables.

3.12 Computational Method for Including Variables

3.12.1 Maximum likelihood Estimation

Logistic regression uses maximum likelihood estimation to compute the coefficients for the logistic regression equation.

Under weak regularity conditions, such as the parameter space having fixed dimension with true value falling in its interior, maximum likelihood (ML) estimation has desirable properties:

- They have large-sample normal distributions
- They are asymptotically consistent; as $n \rightarrow \alpha, \hat{\theta} \xrightarrow{p} \theta$
- They are asymptotically efficient, producing large-sample standard errors not greater than those from other estimation procedures (Agresti, 2002)

Maximum likelihood estimation is one of several alternative approaches that statisticians have developed for estimating the parameters in a mathematical model. Another well-known and popular approach is least squares (LS) estimation. ML estimation and least squares estimation are different approaches that give the same results for classical linear regression compared to least squares, the ML method can be applied in the estimation of complex non linear as well as linear models. In particular, because the logistic model is a non-linear model, ML estimation is the preferred estimation method for logistic regression. ML estimation requires no restrictions of any kind on the characteristics of the independent variable. Thus, when using ML estimation, the independent variables can be nominal, ordinal, and/or interval.

To describe the ML procedure, we introduce the likelihood function (L). Given the data, the likelihood function, L, is the probability of those data, treated as a function of the unknown parameters in one's model, thus can alternatively be denoted as L(θ), where θ denotes the collection of unknown parameters being estimated in the model. In matrix terminology, the collection θ is referred to as a vector; its components are the individual parameters being estimated in the model, denoted here as $\theta_1, \theta_2, \ldots \theta_q$, where q is the number of individual components.

The likelihood function $L(\theta)$ represents the joint probability or likelihood of observing the given data that have been collected. The term "joint probability" means a probability that combines the contributions of all the subjects in the study. That is,

$$L(\theta|x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n)|\theta) = \prod_{i=1}^{n} f\,(x_i|\theta).$$

For convenience, we work with the logarithm of the likelihood function, called the logarithm of the likelihood function, called the log likelihood. Thus $lnL(\theta|x_1, x_2, \ldots, x_i) = \sum_{i=1}^{n} lnf(x_i|\theta)$

There are two alternative ML approaches that can be used to estimate the parameters in a logistic model. These are called the unconditional (L_u) method and the conditional L_c method. In making the choice between unconditional and conditional ML approaches, the researcher needs to consider the number of parameters in the model relative to the total number of subjects under study. In general, unconditional ML estimation is preferred if the number of parameters in the model is small relative to the number of subjects. In contrast, conditional ML estimation is preferred if the number of parameters in the model is large relative to the number of subjects.

As shown below, the ML estimation formula for the unconditional (L_u) approach directly describes the joint probability of the study data as the product of the joint probability for the cases (diseased persons) and the joint probability for the non-cases (non-diseased persons)

$L_u = \prod_{l=1}^{m} p(x_l) \prod_{l=m+1}^{n} [1 - p(x_i)]$; where

M = cases (denoted as $x_1, x_2, \ldots, x_m$); n-m = noncases (denoted as $x_{m+1}, x_{m+2}, \ldots, x_n$), the logistic model, $P(X) = \frac{1}{1+e^{-(\alpha+\sum \beta x_i)}}$

We use these products by assuming that we have independent observations on all subjects. The probability of obtaining the data for the l^{th} case is given $P(x_l)$, where P(X) is the logistic model formula for individual X. the probability of the data for the l^{th} noncase is given by $1 - P(x_l)$.

When the logistic model formula involving the parameters is substituted into the likelihood expression above, the formula shown here is obtained.

$$L_u = \frac{\prod_{l=1}^{n} exp\left(e + \sum_{i=1}^{k} \beta_i x_{il}\right)}{\prod_{l}^{n} \left[\left[1 + \exp\left(\alpha + \sum_{i=1}^{k} \beta_i x_{il}\right)\right]\right]}$$

This expression for the likelihood function L is a function of the unknown parameters and the β_i. The equivalent formula for the conditional (L_c) approach is given by:

$$L_c = \frac{\prod_{l=1}^{n} exp\left(\sum_{i=1}^{k} \beta_i x_{il}\right)}{\prod_{l}^{n}\left[\left[1 + \exp\left(\sum_{i=1}^{k} \beta_i x_{il}\right)\right]\right]}$$

It must be noted however, that the formulae for the likelihood functions for both the unconditional and conditional ML approaches are quite complex mathematically. The formulae are, however built into SPSS. The program does the heavy calculations of forming the likelihood function internally and maximizing this function to obtain the ML solutions.

Once the likelihood function has been determined, the method of maximum likelihood chooses that estimator of the set of unknown parameters; $\theta,$ whch maximizes the likelihood function L(θ). The estimator is denoted as $\hat{\theta} = argmax\hat{l}(e|x_1, x_2, \dots, x_n|), \theta\epsilon 0$ and its components are $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q$.

In general, maximizing the likelihood function $L(\theta)$ is equivalent to maximizing the natural log of $L(\theta)$, which is computationally easier. The components of θ are then found as solutions of equations of partial derivatives: $\frac{\partial lnL(\theta)}{\partial \theta_J} = 0, J = 1,2, \dots, q$

Each equation is stated as the partial derivative of the log of the likelihood function with respect to $\theta_j = 0, where\ \theta_j$ is the jth individual parameter.

If there are q parameters in total, then the above set of equations is a set of q equations with q unknowns. These equations must then be solved iteratively.

3.12.1.1 Likelihood Equations

If we let $x_i = (x_{i1}, \dots, x_{ip})$ deonte setting i of values of p explanatory variables, $i = 1, 2, \dots N,$ then the logistic regression model $P(X) = \frac{1}{1+e^{-(\alpha+\sum \beta_i X_i)}}$, regarding α as a regression parameter with unit coefficient, is $\frac{1}{1+e^{-\left(\sum_{j=1}^{p} \beta_j x_{ij}\right)}}$, When more than one observation occur at a fixed x_i value, it is sufficient to record the number of observations n_i and the number of successes. We then let y_i refer to this success count rather than to an individual binary response. Then $\{Y_1, .., Y_n\}$ are independent binomials with $E(Y_i) = n_i$, where $n_1 + \cdots N_n = n$.

Their joint probability mass function is proportional to the product of N binomial functions.

$$\prod_{i=1}^{N} P(x_i)^{y_i}\,[1 - P(x_i)]^{n_i - y_i} = \left\{\prod_{i=1}^{N} Exp\left[log\left(\frac{P(x_i)}{1-P(x_i)}\right)^{y_i}\right]\right\}\{\prod_{i=1}^{N}[1 - P(x_i)]^{n_i}\}$$

$$\left\{Exp\left[\sum_i y_i\, log\,\frac{P(x_i)}{1-P(x_i)}\right]\right\}\{\prod_{i=1}^{N}[1 - P(x_i)]^{n_i}\}.$$

Therefore, for the logistic regression model $P(x_i) = \frac{1}{1+e^{-\left(\sum_{j=1}^{p}\beta_j X_{ij}\right)}}$, the ith logit is $-\left(\sum_{j=1}^{p}\beta_j X_{ij}\right)$, so the exponential term in the last expression equals

$$exp\left[\sum_j y_j\left(\sum_j \beta_j\, x_{ij}\right)\right] = exp\left[\sum_j\left(\sum_j y_i\, x_{ij}\right)\beta_j\right].$$

Also, since $[1 - P(x_i)] = \left[1 + exp\left(\sum_j \beta_j x_{ij}\right)\right]^{-1}$, the log likelihood equals

$L(\beta) = \sum_j\left(\sum_j y_i\, x_{ij}\right)\beta_j - \sum_j n_i log\left[1 + exp\left(\sum_j \beta_j x_{ij}\right)\right]$. This depends on the binomial counts only through the sufficient statistics $\left\{\sum_j\left(\sum_j y_i\, x_{ij}\right), j = 1, \ldots, p\right\}$.

Likelihood equations result from setting $\partial L(\beta)/\partial(\beta) = 0$. since

$$\frac{\partial L(\beta)}{\partial(\beta)} = \sum_j y_i x_{ij} - \sum_j n_i x_{ij}\,\frac{\exp\left(\sum_j y_i x_{ij}\right)}{1+exp\left(\sum_k \beta_k x_{ik}\right)},$$

The likelihood equations are $\sum_j y_i\, x_{ij} - \sum_j n_i \hat{P}_j x_{ij} = 0, j = 1, \ldots, p,$ where

$\hat{P}_i = exp\left(\sum_j n_i \hat{\beta}_j x_{ij}\right)/\left[1 + exp\left(\sum_j \hat{\beta}_j x_{ij}\right)\right])$ is the ML estimate of $P(x_i)$

The Equations are nonlinear and require iterative solution.

Let X denote the $N \times p$ matrix of values of x_{ij}. The likelihood equations have form

$X'y = X'\hat{\mu}, where\ \mu = n_i \hat{p}_i$

3.12.1.2 Newton Raphson Iterative Method

Newton-Raphson method is an iterative method for solving nonlinear equations whose solutions determine the point at which a function takes its maximum. It begins with an initial guess for the solution. It obtains a second guess by a second degree polynomial and then finding the location of the polynomial's maximum value. It then approximates the function in a neighbourhood of the second guess by another second-degree polynomial, and the third guess is a location of its maximum. In this manner, the method generates a sequence of guesses. These converge to the location of the maximum when the function is suitable and/or the initial guess is good. In more detail, here is how Newton-Raphson determines the value $\hat{\beta}$ at which a function $L(\beta)$ is maximized.

Let $\mu' = \left(\partial L(\beta)/\partial\beta_{1,}\ \partial L(\beta)/\partial\beta_{2,} \ldots, \partial L(\beta)/\partial\beta_{j,}\right)$

Let H denote the matrix having entries $h_{ab} = \partial^2 L(\beta)/\partial\beta_a \partial\beta_{b,}$ called the Hessian matrix. Let $H^{(t)}\ and\ H^{(t)}\ be\ u\ and\ H$ evaluated at $\beta^{(t)}$ then the guess t for $\widehat{\beta}.\ step\ t$ in the iterative process.

(t = 1 , 2...) approximates L(β) $near\ \beta^{(t)}$by the terms up to the second order in its taylor series expansion

L(β) $= L(\beta^t) + u^{(t)}(\beta - \beta^t) + \left(\frac{1}{2}\right)\left(\beta - \beta^{(t)}\right)H^{(t)}\left(\beta - \beta^{(t)}\right).$

Solving $\frac{\partial L(\beta)}{\partial \beta} = u^t + H^{(t)}\left(\beta - \beta^{(t)}\right) = 0\ for\ \beta$ yields the next guess.

That guess can be expressed as

$\beta^{(t+1)} = \beta^{(t)} - \left(H^{(t)}\right) - 1_{u^{(t)}}$, assuming that $H^{(t)}$ is nonsingular. Iteration proceeds until changes in $L\left(\beta^{(t)}\right)$ between successive cycles are sufficiently small. The ML estimator is the limit of $\beta^{(t)}\ as\ t \longrightarrow \infty$ (walker and Duncan, 1967)

For many models $L(\left(\hat{\theta}\right)$ has concave shape and θ is the point at which $\frac{\partial InL(\theta)}{\partial \theta_j} = 0$

The overall measure of how well the model fits is given by the likelihood value, which is similar to the residual or error sum of squares values for multiple regressions. A model that fits the data well will have a small likelihood value. A perfect model would have a likelihood value of zero. Maximum- likelihood estimation is an interactive procedure that successively tries to work to get closer and closer to the correct answer.

3.13 Statistical Inferences Using Maximum Likelihood Techniques

Once ML estimates have been obtained, these estimates can now be used to make statistical inferences concerning the exposure-disease relationship under study.

The maximum likelihood value, $L\left(\hat{\theta}\right)$ which is the numerical value of the likelihood function l when the ML estimates are substituted for their corresponding parameter values can be used to illustrate how statistical inferences are made.

We consider the following three hypothetical models, each written in logit form:

Model 1: logit $P_1(X) = \alpha + \beta_1 X_1 + \beta_2 X_2$

Model 2: logit $P_2(X) = \alpha + \beta_1 X_1 + \beta_2 X_2 B_3 X_3$

Model 3: logit $P_3(X) = \alpha + \beta_1 X_1 + \beta_2 X_2 B_3 X_3 + \beta_4 X_1 X_3 + \beta_5 X_2 X_3$

Model 1 involves two variables $X_1 and\ X_{2.}$ model 2 contains these same two variables and a third variable X_3. Model 3 contains the same three X's as in model 2 plus two additional variables, which are the product terms X_1, X_3 and $X_2, X_3. Let\ \widehat{L_{1,}}$ $\widehat{L_2}$ and $\widehat{L_3}$ donate the maximized likelihood values based on fitted models 1,2 and 3 respectively.

The fitting may be done either by unconditional or conditional methods, depending on which method is more appropriate for the model and data set being considered.

Since a model fits the data better when it has more parameters, it follows that $\widehat{L_1} \le \widehat{L_2} \le \widehat{L_3}$

This relationship among the $\widehat{Ls}$ is similar to the property in classical multiple linear regression analyses that the more parameters a model has, the higher is the R square

(R^2) statistic for the model. In other words, the maximized likelihood value $\hat{L}$ is similar to R square, in that the higher the $\hat{L}$, the better the fit.

It follows from algebra that if $\widehat{L_1} \le \widehat{L_2} \le \widehat{L_3}$ then the same inequality relationship holds for the natural logarithms of these $\widehat{Ls}, that\ is, ln\widehat{L_1} \le ln\widehat{L_2} \le ln\widehat{L_3}$.

However, if we multiply each log of $\hat{L}$ by -2, then the inequalities switch around so that $-2ln\widehat{L_3} \le -2ln\widehat{L_2} \le -2ln\widehat{L_1}$.

The statistic $-2ln\widehat{L_1}$ is called the log likelihood statistic for model 1, similarly, the other statistic are the log likelihood statistic for their respective models. These statistics are important because they are used to test hypotheses about parameter in the model using what is called a likelihood ratio test.

3.14 The likelihood Ratio Test

Statisticians have shown that the difference between log likelihood statistics for two models $\{-2lnL_1 - (-2lnL_2)\}$, one of which is a special case of the other, has an approximate chi-square distribution in large samples. Such a test statistic is called a likelihood ratio statistic (LR) the degrees of freedom (df) for this chi-square test are equal to the difference between the number of parameters in the two models.

In general, the likelihood ratio statistic requires the identification of two models to be compared; of which one is a special case of the other. The large model called the full model and the smaller model called the reduced model; that is obtained by setting certain parameter in the full model equal to zero. Given for example, two models:

Model 1: logit $P_1(X) = \alpha + \beta_1 X_1 + \beta_2 X_2$

Model 2: logit $P_2(X) = \alpha + \beta_1 X_1 + \beta_2 X_2 B_3 X_3$

Model 1 is a special case of model 2 by settings $B_3 = 0$

The set of parameters is in the full model; that is set equal to zero specify null hypothesis being tested.

H_0: parameter in full model = 0

df= number of parameters set equal to zero correspondingly, the degrees of freedom for the likelihood ratio test are equal to the number of parameters in the large model that must be set equal to zero to obtain the small model.

3.15 How the LR Test Work

Algebraically, LR= $-2lnL_1 - (-2lnL_2)$= -2ln $\frac{\widehat{L_1}}{\widehat{L_2}}$

If the addition variable X_3, makes an extremely large contribution to the risk of disease over that already contributed by X_1 and X_2 then it follows that the maximized likelihood value L_2 is much larger than the maximized likelihood value L_1

If L_2 is much larger than L_1, then $\frac{L_1}{L_2} \approx 0$

Now the natural log of any fraction between 0 and 1 is a negative number. As this fraction approaches 0 ln $\frac{\widehat{L_1}}{\widehat{L_2}} \approx \infty$

Thus, the LR for a highly significant X_3 variable is large and positive and approaches ∞. In contrast, if the additional variable, X_3 makes no contribution whatsoever to the risk of disease over and above that contributed X_1 and X_2, $this\ would\ mean\ that$ $\widehat{L_1} \approx \widehat{L_2}$.

$\Rightarrow$ LR = $\frac{\widehat{L_1}}{\widehat{L_2}} \approx 1$

$\Rightarrow LR \approx -2\ln(1) = -2 \times 0 = 0$

Thus, the LR for a highly non-significant X_3 variable is approximately 0. In summary, the likelihood ratio statistic, regardless of which two models are being compared, yields a value that lies between 0, when there is extreme non-significance, and ∞, when there is extreme significance ($0 \leq LR \leq \infty$). The LR test does this by comparing the log likelihood of the two models, if this difference is statistically significant, then the less restrictive model (the one with more variables) is said to fit the data significantly better than the more restrictive model.

3.16 Wald Test

A wald test is used to test the statistical significance of each coefficient (β) in the model. A wald test calculates a Z statistic, which is: z = $\frac{\hat{\beta}}{sE}$

This z value is then squared, yielding a Wald statistic with a chi- square distribution. However, several authors have identified problems with the use of the Wald statistic. Menard (1995) warns that for large coefficients, standard error is inflated, lowering the Wald statistic (chi-square) value. Agresti (1996) states that the likelihood-ratio test is more reliable for small sample sizes than the Wald test

3.17 Overall Model Fit

The null model -2 log likelihood is given by -2* in (L_0) where L_0 is the likelihood of obtaining the observations if the independent variables had no effect on the outcome.

The full model -2 Log likelihood is given by -2* ln (L) where L is the likelihood of obtaining the observations with all independent variables incorporated in the model.

The difference of these two yields a Chi-square statistic which is a measure of how well the independent variables affect the outcome or dependent variable.

If the P-value for the overall model fit statistic is less than the conventional 0.05 then there is evidence that at least one of the independent variables contributes to the prediction of the outcome.

3.18 Interpreting Parameters in Logistic Regression

3.18.1 Odds

For a probability D of success, the odds are defined to be $\frac{p(x)}{1-p(x)}$, the odds are nonnegative, with $\frac{p(x)}{1-p(x)} > 1$ when a success is more likely than a failure. The quantity $\frac{p(x)}{1-p(x)}$, whose log value gives the logit, describes the odd for developing the disease for a person with independent variables specified by x_i where P(x) donates the probability that an event will occur and 1-P(x), the probability that the same event will not occur.

The logit form of the logistic model, shown again here, gives an expression for the log odds of developing the disease for an individual with a specific set of X's.

Logit (x) = $ln_e\left[\frac{p(x)}{1-p(x)}\right]$ = log odds for individual X = $\alpha + \sum \beta ixi$

Where $\frac{p(x)}{1-p(x)}$ describes risk in logistic model for individual X

The ratio of two odds is called the odds ratio

3.18.2 Odd Ratio

3.18.2.1 Definition

The odds ratio for a predictor is defined as the relative amount by which the odd of the outcome increase (OR greater than 1.0) or decrease (OR less than 1.0) when the value of the predictor variable is increased by 1.0 units. In other words, (odds for PV+ 1)/(odds for PV), where PV is the value of the predictor variable.

Any odds ratio, by definition, is a ratio of two odds, written for example as $\frac{p(x_2)}{1-p(x_2)} / \frac{p(x_1)}{1-p(x_1)}$ (odd$s_2/odds_1$), in which the subscript indicates two individuals or two groups of individuals being compared. For joint distribution with cell probabilities (D_{ij}), the equivalent definition for the odds in row *i* is D_{ij}/D_{i2} = 1,2. Then the odds ratio

$$\frac{D_{11}/D_{12}}{D_{21}/D_{22}} = \frac{D_{11}D_{22}}{D_{12}D_{11}}$$

More generally, when we describe an odds ratio, the two groups being compared can be defined in terms of X_i which donates a general collection of X variables, from 1 to k.

$X_2 = (X_{21,}X_{22,} \dots X_{2k})$

X_1 =$(X_{11}, X_{12}, \dots X_{1k})$

Let X_2 donate the collection of X's that specify group 2 and let X_1 donate the collection of X's that specify group 1.

Notationally, to distinguish the two groups X_2 and X_1 in an odd ratio, we write

ORx_2x_1= $\frac{odds\ for\ X_2}{odds\ for\ X_1}$

We now apply the logistic model to this expression to obtain a general odds ratio formula involving the logistic model parameters.

3.18.2.2 Derivation of OR Formula
Given a logistic model of the general form

p(x)= $\frac{1}{1+e^{-(\alpha\Sigma\beta x_i)}}$, we write the odds for group2 as $\frac{p(x_2)}{1-p^{(x_2)}} = \frac{\frac{1}{1+e^{-(\alpha+\Sigma\beta_i x_{2i})}}}{1-\frac{1}{1+e^{-(\alpha\Sigma\beta_i x_{2i})}}} = e^{(\alpha+\Sigma\beta_i x2i)}$

And the odds for group1 as, $\frac{p(x_2)}{1-p(x_2)} = \frac{\frac{1}{1+e^{-(\alpha+\Sigma\beta_i x_{1i})}}}{1-\frac{1}{1+e^{-(\alpha\Sigma\beta_i x_{1i})}}} = e^{(\alpha+\Sigma\beta_i x1i)}$

To get an odds ratio, we then divide the first odds by the second odds. The result is an expression for the odds ratio written in terms of the two risks P(X_2) and risks P(X_1). this is for

P(x) = $\frac{1}{1+e^{-(\alpha+\Sigma\beta_{i}x_i)}},$ $OR\ \ x_2, x_1 = \frac{e^{(a+\Sigma\beta_{i}x_{2i})}}{e^{(\alpha+\Sigma\beta_{i}x_{1i})}}$

$$=e^{\Sigma ki=1\ \beta_{i(x_{2i}-x_{1i})}}$$

We thus have a general exponential formula for the logistic model comparing any two groups of individuals, as specified in terms of $X_2\ and\ X_1$. We can give an equivalent alternative to our OR formula by using the algebraic rule that say that exponential of a sum is the same as the product of the exponential of each term in the sum.

That is $e^{\Sigma ki} = 1\beta_i(x_{21} - x_{11})e^{\beta_3}(x_{22} - x_{12}) \ldots e^{\beta_k}(x_{2k} - x_{1k})$

We alternatively write this expression using the product symbol Π, where Π is a mathematical notation which denotes the product of a collection of terms. Thus, using algebraic theory, we obtain the alternative formula for OR as:

$$\Pi_{i}^{k}{=}_1 e^{\beta_i}(x_{21} - x_{11})e^{\beta_3}(x_{22} - x_{12})..e^{\beta_k}(x_{2k} - x_{1k})$$

Thus, the product formula for OR tells us that, when the logistic model is used, the contribution of the variables to the odds ratio is multiplicative.

3.18.2.3 Properties of The odd Ratio

The odds ratio can be equal to any nonnegative number. The condition $\frac{p(x_2)}{1-p(x_2)} = \frac{p(x_1)}{1-p(x_1)}$ and hence (when all cell probabilities are positive) θ= 1 corresponds to independence of X and Y. when 1< $\theta < \infty$, subjects in row 1 are more likely to have a success than subjects in row 2, that is $D_1 > D_{2.}$ $when$ $0 < \theta < 1, D_1 < D_2$ .When one cell has zero probability, θ equals to 0 or ∞. Values of θ farther from 1 in a given direction represent stronger association. Two values represent the same association, but in opposite directions, when one is the inverse of the other. For inference, we use log θ. Independence corresponds to log $\theta = 0$. the log odds ratio is symmetric about this value-reversal of rows or of columns results in a change in its sign. Two values for log θ that are the same except for sign represent the same strength of association.

3.18.3 Regression Coefficients ($\boldsymbol{\beta}_i$)

The regression coefficients are the coefficients $b_0, b_1, b_2, \ldots b_k$ of the regression equation:

Logit (p) = $b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_kX_k$

An independent variable with a regression coefficient not significant from 0 ($P > 0.05$) can be removed from the regression model. If $P < 0.05$, then the variable contributes significantly to the prediction of the outcome variable.

The logistic regression coefficients show the change (increase when $b_i > 0, decrease\ when\ b_i < o$) in the predicted log odds of having the characteristic of interest for a one- unit change in the independent variables.

Given the model logit P(x) = $\alpha+\sum\beta_iX_i$ refers to the effect of X_i on the log odds the D = 1 controlling the other X_j. This provides a basic interpretation for the magnitude of β_i. For instance e^{β_i} is the multiplicative effect on the odds of s_1 – unit increase in x_i, at fixed level of other x_j. The sign of β_i determines whether P(X) is increasing or decresing as x increases. The rate of climb or decent increases as IβI increases; as $\beta \rightarrow 0$, the curve straightens a horizontal straight line when $\beta = 0$, is independent of X. for quantitative X_i with $\beta > 0$,the curve for P(X) approaches 1 at the same rate that it approaches 0.

3.18.4 The Concept (∝)

As a simple example, consider what the logit becomes when all the X's are 0. To compute this, we need to work with the mathematical formula, which involves the unknown parameters and the X's.

If we plug in 0 for all the X's in the formula, we find that the logit of P(X) reduces simply to α. Because we have already seen that any logit can be described in terms of an odds, we can interpret this result to give some meaning to the parameter α.

One interpretation is that α gives the log of the background, or baseline, odds. This interpretation for α is more appealing: to describe it as the log of the background, or baseline, odds. By background odds, we mean the odds that would result for a logistic model without any X's at all. The form of such a model is $P(x) = \frac{1}{1+e^{-\alpha}}$. We might be interested in this model to obtain a baseline risk or odds estimate that ignores all possible predictor variables. Such an estimate can serve as a starting point for comparing other estimates of risk or odds when one or more X's are considered.

3.18.5 The P-value

The p-value is the probability of seeing result as extreme as the one we are getting in a collection of random data in which the variable had no effect. A P of 5% or less is the generally accepted point at which to reject the null hypothesis. With a p - value of 0.05 there is only a 5% chance that results we are seeing would have come up in a random distribution, so we can say with a 95% probability of being correct that the variable is having some effect, assuming our model is specified correctly. When model has a high p-value (p-value >0.05) there is a very good chance that the model is not significant and should not be used.

CHAPTER FOUR

DATA COLLECTION AND ANALYSIS

4.1 Introduction

This Chapter, first of all, presents the exploration of the data for important features using summary statistics and histograms. The logistic model would be used to determine the risk factors affecting stillbirth in the Hohoe Municipal hospital. Also, we shall examine the correlations of the coefficients of the independent variables.

4.2 Descriptive Analysis

Refer to graphical representation in Appendix B for most of the descriptive analysis done here.

It was observed that out of 250 women interviewed about 12.4% of them had stillbirths whereas about 87.6% had live births. Thus, there were more children born alive compared to those who were stillbirths.

Moreover, about 14 (5.6%) of the women considered for the study were less than 18 years and about 213 (85.2%) of them were between 18 years and 34 years. Only 23 (9.2%) of the women were 35 years and above. It was noticed that 119 (47.6%) of the women lived in urban community as against 131 (52.4%) of those who lived in rural communities. This shows that there were about 11 more women in rural community than urban. Again, it was seen that, about 133 (53.2%) of the respondents were females whilst about 117 (46.8%) of children were males. This shows that there were about 16 more female babies compared to their male counterparts.

Further, it was observed that, about 152 (60.8%) of the respondents were diagnosed not to have encountered obstetric abnormalities during their pregnancy whilst about 98 (39.2%) of the women were diagnosed with obstetric abnormalities during pregnancy at the Hohoe Municipal hospital.

Also, out of the 250 respondents about 155(62%) of them were not taking alcoholic drinks whilst 95 (38%) confirmed they took alcoholic drinks. Meanwhile, about 188(75.2%) of the respondents interviewed said they were not practicing self-medication during pregnancy and about 62(24.8%) of them confirmed they were practicing self-medication during their pregnancy.

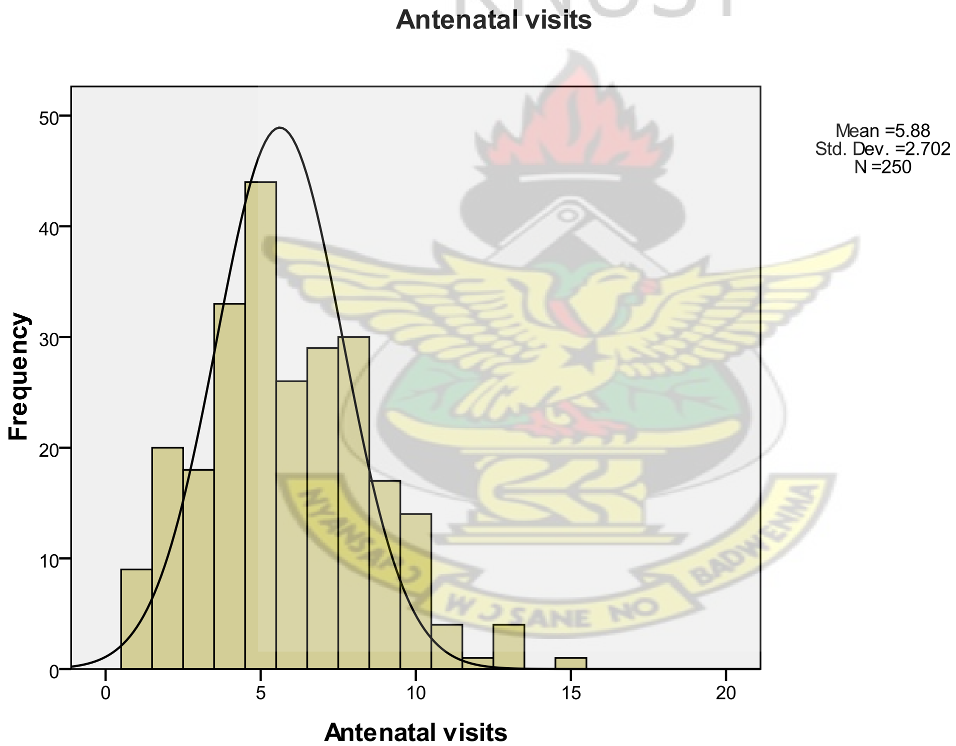


Fig. 4.1 The distribution of antenatal visits by pregnant women

From Fig. 4.1, it shows that the most frequent antenatal visits were 5 days whilst the mean pregnant women antenatal visits was 5.88 days to the Hohoe Municipal hospital with a standard deviation of about 2.702. The antenatal distribution of pregnant women antenatal visits as observed from the figure is fairly normal.

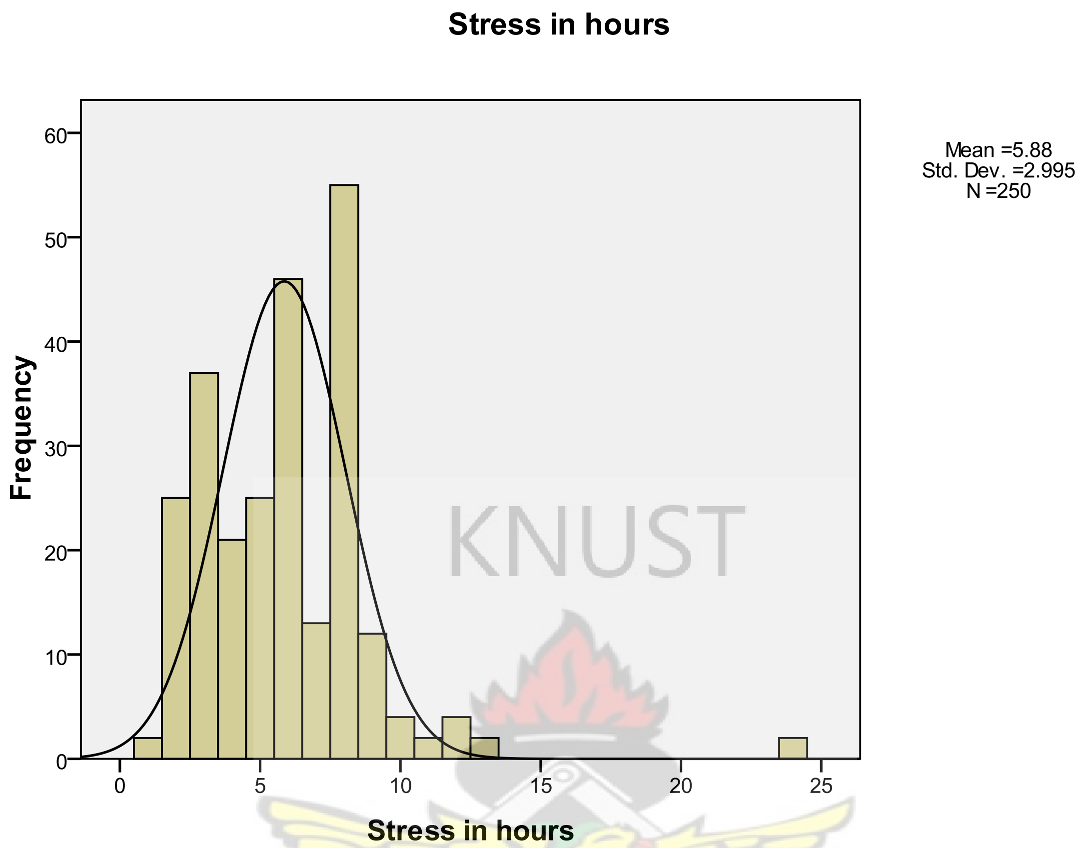


Fig. 4.2 Daily hours worked by pregnant women

Figure 4.2 above indicates that, the most common hours worked by the women interviewed was 8 hours and the mean hours worked during pregnancy by the same women was 5.88 hours with a standard deviation of about 2.995.

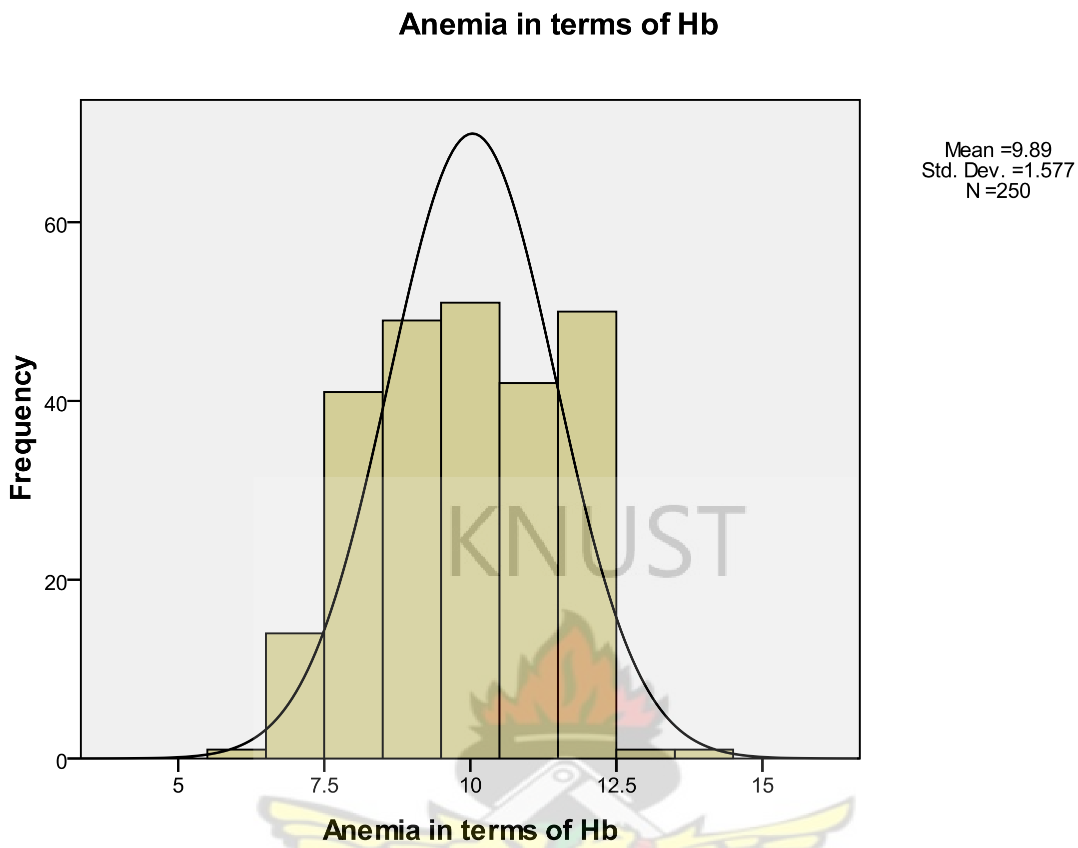


Fig. 4.3 Anaemia status of pregnant women in Hb

Further, Fig.4.3 shows that, the most common Haemoglobin measured to assess the anaemia status of pregnant women was 10glb and the mean of these women was 9.89glb with a standard deviation of about 1.577. The anaemia status of the respondents given in terms of Haemoglobin (Hb) distribution was fairly normal.

4.3 Inferential Analysis of Data

Table 4.1 Collinearity Diagnostic Test

Variables	Tolerance
Age group	0.932
Place of residence	0.844
Antenatal visits	0.882
Weight of baby	0.803
Sex of baby	0.941
Obstetric problem	0.837
Alcoholic drink	0.922
Self-medication	0.800
Stress in hours	0.918
Anaemia in terms of Hb	0.972

From Table 4.1, it can be seen that the tolerance values for each of the independent variables age group, place of residence, antenatal visits, weight of baby, sex of baby, obstetric problem, alcoholic drink, self-medication, stress in hours and anaemia in terms of Hb were each less than the reference value 1. This means that there is no interaction between the independent variables. Therefore, all the independent variables were used for modelling stillbirth at the Hohoe Municipal hospital from January – June 2011.

Now, it is important to note that, the decision on which logit coefficient is significant to the model, can be determined by several techniques in logistic regression model building. This techniques involves using either the Wald test or the log likelihood (-2logL) or by comparing the p–value with the significance level (5%). Significance is established if p–value is ≤ 0.05. For this study, the p-value was used with significance level of $\alpha = 0.05$ for the analysis.

Table 4.2 Logistic Regression Predicting Likelihood of Stillbirth

Variable	β	S.E.	Wald	P-value	Odds Ratio	95% C.I. for Odds Ratio	
						Lower	Upper
AGE	16.985	8.808	.000	.999	2.387	.556	6.836
RESIDENCE	.163	.596	.075	.044	1.178	.366	3.787
ANTENATAL	.110	.110	.992	.319	1.116	.899	1.385
WEIGHT	-.504	.378	1.776	.183	.604	.288	1.268
SEX	.488	.531	.847	.357	1.630	.576	4.610
OBSTETRIC	2.145	.680	9.955	.002	8.540	2.253	32.365
ALCOHOL	.725	.530	1.867	.012	2.064	.730	5.836
MEDICATION	.255	.581	.193	.032	1.291	.413	4.031
HOURS	.101	.100	1.024	.312	1.106	.910	1.346
ANAEMIA	-.304	.173	3.082	.079	.738	.525	1.036
Constant	-18.506	8.808	.000	.999	.000		

From Table 4.2 above, the logistic model was obtained as follows

Logit (P(y=1)) = -18.506+0.163Resident+2.145Obstetric+0.725Alcohol+0.255Medication

(4.1)

Hypothesis testing using in Table 4.2 above

H_0: $\beta_J = 0$

H_1: $\beta_J \neq 0$

Where j= each coefficient of the independent variables in Table 4.2

Explanation of variables in the model

It can be noted from Table 4.2 that, the independent variables; place of residence, obstetric problems, alcoholic drinks and self-medication with the p-values 0.044, 0.002, 0.012 and 0.032 respectively are each less than $\alpha = 0.05$. Therefore we reject the null hypothesis and conclude that the coefficients of these variables are each not equal to zero at 95% confidence interval. This means that these predictors are each important to be included in the final model 4.1. Based on this reason, therefore, we can say that these predictors are relevant predicting stillbirth at Hohoe Municipal hospital from January – June 2011.

Explanation of variables not in the model

From Table 4.2 below, it is worth noting that, the independent variables; age group, antennal visits, weight of baby, sex of baby, hours of work and anaemia were dropped from the model. Since the p – values 0.999, 0.319, 0.183, 0.357, .312 and 0.079 were each greater than $\alpha =$ 0.05, so we fail to reject the null hypothesis and conclude that the coefficients of these variables are each equal to zero. This shows that these independent variables were not important to be included in the model. Hence the independent variables age group, antenatal visits, weight of baby, sex of baby, hours of work and anaemia were not relevant in predicting stillbirth at Hohoe Municipal hospital from January – June 2011.

Odds ratios Interpretations

As in Table 4.2, the strongest predictor of stillbirth was obstetric problem, recording an odds ratio of 8.540 (95 % C.I. =2.253 – 32.365). This indicated that pregnant women who had obstetric problem were over 8 times as likely to estimate stillbirth as those who did not experience obstetric problem, controlling for all other factors in the model. The odds ratio 1.178 (95% C.I. = 0.366 – 3.787) for place of residence indicates that women from rural communities had stillbirths compared to their counterparts from urban communities, holding other factors in the model constant.

Again, the odds ratio with respect to alcohol was 2.064 (95% C.I. = 0. 730 – 5.836) meaning that more of the stillbirth was estimated by pregnant women who take alcoholic drinks compared to those who were not taking alcoholic drinks, holding other factors constant. Moreover, self-medication had odds ratio of 1.291 (95% C.I. = 0.413 – 4.031). This implies that, stillbirth was predicted more by self-medicated pregnant women than those women who did not practice self-medication during pregnancy, if other factors remain constant.

Table 4.3 Assessing Model Fit by Hosmer and Lemeshow Test

Chi-square	Df	P-value
6.948	8	0.542

H_0: The model fits the data

H_1: The model does not fit the data

From Table 4.3, since the p – value, 0.542, is greater than the significance level, $\alpha = 0.05$, we fail to reject the null hypothesis (H_0) and conclude that there is enough evidence to show that the hypothesized model fits the data set used in predicting stillbirth at Hohoe Municipal hospital from January – June 2011.

Table 4.4 Model Summary

-2 Log likelihood	Cox & Snell R-Square	Nagelkerke R-Square
566.036	0.281	0.635

It is worth noting from Table 4.4 that, between 28.1% and 63.5% of the variance in predicting stillbirth was explained by the independent variables; place of residence, antenatal visits, alcohol intake and self-medication. However, it was shown from the classification table (see appendix B) that, about 87.6% could be predicted alive whilst about 12.4% can be predicted dead. It is worth noting that, overall, about 95.8% of the cases were correctly classified.

Table 4.5 Correlation Matrix of coefficients of independent variables

	Age	Res	Ante	Wei	Sex	Obste	Alco	Med	Hrs	Ane
Age	1.000									
Re	0.000	1.000								
Ant	0.000	0.119	1.000							
Wei	0.000	0.121	**-0.318**	1.000						
Sex	0.000	-0.006	-0.065	-0.021	1.000					
Obst	0.000	0.026	0.020	0.068	0.128	1.000				
Alco	0.000	0.013	0.030	0.003	-0.082	0.192	1.000			
Med	0.000	0.052	0.266	-0.031	0.176	-0.081	-0.068	1.000		
Hrs	0.000	0.219	0.140	-0.120	0.148	0.161	0.143	**0.326**	1.000	
Ane	0.000	-0.038	-0.055	0.132	-0.013	-0.264	-0.206	-0.016	-0.171	1.000

From Table 4.5, it can be noted that the largest correlation coefficient is the coefficient between number of hours worked and self-medication. This coefficient equals 0.326 which means that self-medication of the pregnant women interviewed increases as hours worked increases. However, the largest negative correlation coefficient was between weight of baby and antenatal visits, implying that as the number of antenatal visits increases the weight of baby decreases. Meanwhile, since there is no correlation coefficient greater than 0.9 there does not seem to be severe multicollinearity between the independent variables.

Table 4.6 Chi-square Test of Association between Sex and Stillbirth

	Value	Df	P-value
Pearson Chi-Square	2.984	1	.084
Likelihood Ratio	2.990	1	.084
Linear-by-Linear Association	2.972	1	.085
N of Valid Cases	250		

The null and the alternative hypothesis for assessing the association between sex and stillbirth

H_0: Sex of baby is not associated with stillbirth

H_1: Sex of baby is associated with stillbirth

From Table 4.6 above, since the p-value = 0.084 is greater than the significance level, α = 0.05, we fail to reject the null hypothesis and conclude that stillbirth is independent of sex of a baby at 95% confidence level.

CHAPTER FIVE

FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

5.1 INTRODUCTION

This chapter seeks to find the risk factors affecting stillbirth at the Hohoe Municipal hospital in the Volta Region. This Chapter summarizes the findings and conclusions drawn from the analysis and some suggested recommendations.

5.2 FINDINGS

From the analysis done in Chapter four, some findings were made. About 60.8% of the respondents were diagnosed not to have encountered obstetric abnormalities during their pregnancy whilst about 39.2% of the women were diagnosed with obstetric abnormalities during pregnancy at the Hohoe Municipal hospital. Also, 62% of them were not taking alcoholic drinks whilst 38% confirmed those taking alcoholic drinks. Meanwhile, about 75.2% of the respondents interviewed said they were not practicing self-medication during pregnancy and about 24.8% of them confirmed they were practicing self-medication during their pregnancy. It was found that the most frequent antenatal visits were 5 days whilst the mean pregnant women antenatal visits were 5.88 days to the Hohoe Municipal hospital.

The independent variables; place of residence, obstetric problems, alcoholic drinks and self-medication were relevant predictors of stillbirth whilst the independent variables; age group, antennal visits, weight of baby, sex of baby, hours of work and anaemia were dropped from the model.

The final model was obtained as follows

$$\text{Logit}\ (P(y=1)) = -18.506+0.163\text{Resident}+2.145\text{Obstetric}+0.725\text{Alcohol}+0.255\text{Medication}$$

Meanwhile, the strongest predictor of stillbirth was obstetric problem, recording an odds ratio of 8.540 (95 % C.I. =2.253 – 32.365), which indicated that pregnant women who had obstetric problem were over 8 times as likely to estimate stillbirth as those who did not experience obstetric problem, controlling for all other factors in the model.

5.3 CONCLUSIONS

It was concluded that, among the ten risk factors considered, the risk factors; place of residence, obstetric problems, alcoholic drinks and self-medication were relevant predictors of stillbirth whilst age group, antennal visits, weight of baby, sex of baby, hours of work and anaemia were not important in predicting still births at the Hohoe Municipal hospital. Again, some of the risk factors were correlated with the largest correlation coefficient being the coefficient between number of hours worked and self-medication. This coefficient equals 0.326 which means that self-medication of pregnant women interviewed increases as hours worked increases. Finally, it was concluded that stillbirth is independent of sex of a baby at 95% confidence level.

5.4 RECOMMENDATIONS

The following recommendations were made:

(i) The Ghana Health Service (GHS) should sensitize pregnant women on the dangers of alcoholic drinks to the foetus, so as to avoid stillbirths.

(ii) The Ministry of Health (MoH) could liaise with the information service to make public to the rural folks about the need to be visiting hospitals or clinics for professional attention and the effects of self-medication so as to reduce if not eliminate still birth.

(iii) Pregnant women should be educated to be doing minimal exercise rather than working more hours at the expense of their pregnancy.

(iv) Further research could be done by increasing the study area and for that matter the sample size to examine the relevance of some of the medical risk factors which are yet to be proven by literature.

REFERENCES

Bayaga, A. (2010). Multinomial Logistic Regression: Usage and Application in Risk Analysis, JAQM, Vol. 5, number 2.

Belsten, L.M. (2008). The Silent Womb, State University f New York at Stony Brook.

Cripe, S. M. (2007). Risk Factors Associated with Stillbirth in Thai Nguyen Province, Vietnam, Department of Epidemiology, University of Washington, Oxford University Press.

Derr, R. E. (2009). Performing Exact Logistic Regression with the SAS system, SAS Institute Inc., Cary, NC

Edmond, K.M. (2011). Causes of Stillbirth, Neonatal death and Early Childhood death in rural Zambia by verbal autopsy assessments, Blackwell Publishing Ltd, 7th April, 2011.

Flom, P. L. (1997). Ordinal Logistic Regression in Medical Research, Journal of Royal College of Physicans of London, Vol. 31, Issue 5 pp 546-551.

Haberman, S.J. and Sinharay, S. (2010). The Application of the Cummulative Logistic Regression Model to Automated Essay Scoring.

Hale, M. (2001). Logistic Regression for Geologically Constrained Mapping of Gold Potential, Baguio District, Philippines, Vol. 10; No 3, pp 165-175; Doi: 10.2113/0100165

Hosmer, D. W. and Lemshow, S. (2000). Applied Logistic Regression, USA, John Wiley and Sons, Inc, 111 River Street, Hoboken, NJ 07030,(201) 748-601

Hosmer, D.W. and Lemshow, S. (2000). An introduction to Categorical Data Analysis, New York. http://www.statsdirect.com/help/referecces/refs.htm

Hossain, N. and Khan, N. H. (2009). Obstetric Causes of Stillbirth at Low Socio-Economic Settings, Journal of the Pakistan Medical Association (2009) Vol. 59, Issue 11 pp 744-747, www.ncbi.nlm.nih.gov

Hung, H. and Chen-Chien, W. (2011). Matrix Variate Logistic Regression Analysis, Cornell

Kayri, M. and Cokluk, O. (2010). Using Multinomial Logistic Regression Analysis in Artificial Neural Network, Ozean Journal of Applied Sciences 3(2), 2010, ISSN 1943-2429.

Keating, K. A. and Cherry, S. (2004). Use and interpretation of Logistic Regression in Habitat Selection Studies, Journal of Wildlife Management, October 2004, Vol. 68 issue 4, pp 774-789

Kesmodel, U. (2001). Moderate Alcohol intake during Pregnancy and the Risk of Stillbirth and Death in the first year of life, Am J, Epidemiol (2002) 155(4): pp 305-312

Kleinbaum, D. G. and Klein, M. (1994). Logistic Regression methods, Springer, Atlanta, GA, Birmingham.

Konglai, Z.H.U. and Jingjing, L.I. (2010). Management Science and Engineering Vol. 4, No 4, 2010, pp 24-32; www.cscanada.org

Kristula, (1992). Preventive Vertinary Medicine, Vol. 14, Issues 1-2, Oct 1992, pp 57-68

Kvinnokliniken in Hulthén-Varli, P. I (2010). http://www.ncbi.nlm.nih.gov/Pubmed. Obstet Gynecol 2009 Oct, 114(4): 901-914 Erratum in Obstet Gynecol. 2010 Jan; 115(1):191

Lassey, A.T. and Obed, S.A. (2002). www.ncbi.nlm.nih.gov/Pubmed, Hepatitis C Virus Sero-prevalence Among Mothers Delivering at the Kole-Bu Teaching Hospital, Ghana, Department of Obstetrics and Gynaecology, University of Ghana, Accra.

Lee, J. (1985). An Insight on the Use of Multiple Logistic Regression Analysis to Estimate Association between Risk Factor and Disease Occurrence; Oxford Journals, Vol. 15, Issue1, pp 22-29

Marbury, M.C. (1983). The Association of Alcohol Consumption with Outcome of Pregnancy, American Journal of Public Health 73, No. 10 pp 1165 -1168, Oct 1983

Reddy, U. M. (2006). Maternal Age and Risk of Stillbirth, Dept of Health and Human Services, Bethesda, MD, USA.

Reddy, U.M (2011). Pregnancy Risk Factors for Antepatum Stillbirth in the USA in Obstetrical & Gynaecological Survey: Feb 2011 - Vol. 66, issue 2, pp 69-70.

Say, L. (2006). Journal on Reproductive Health, Vol. 3,Issue 1, 1st October, 2006.

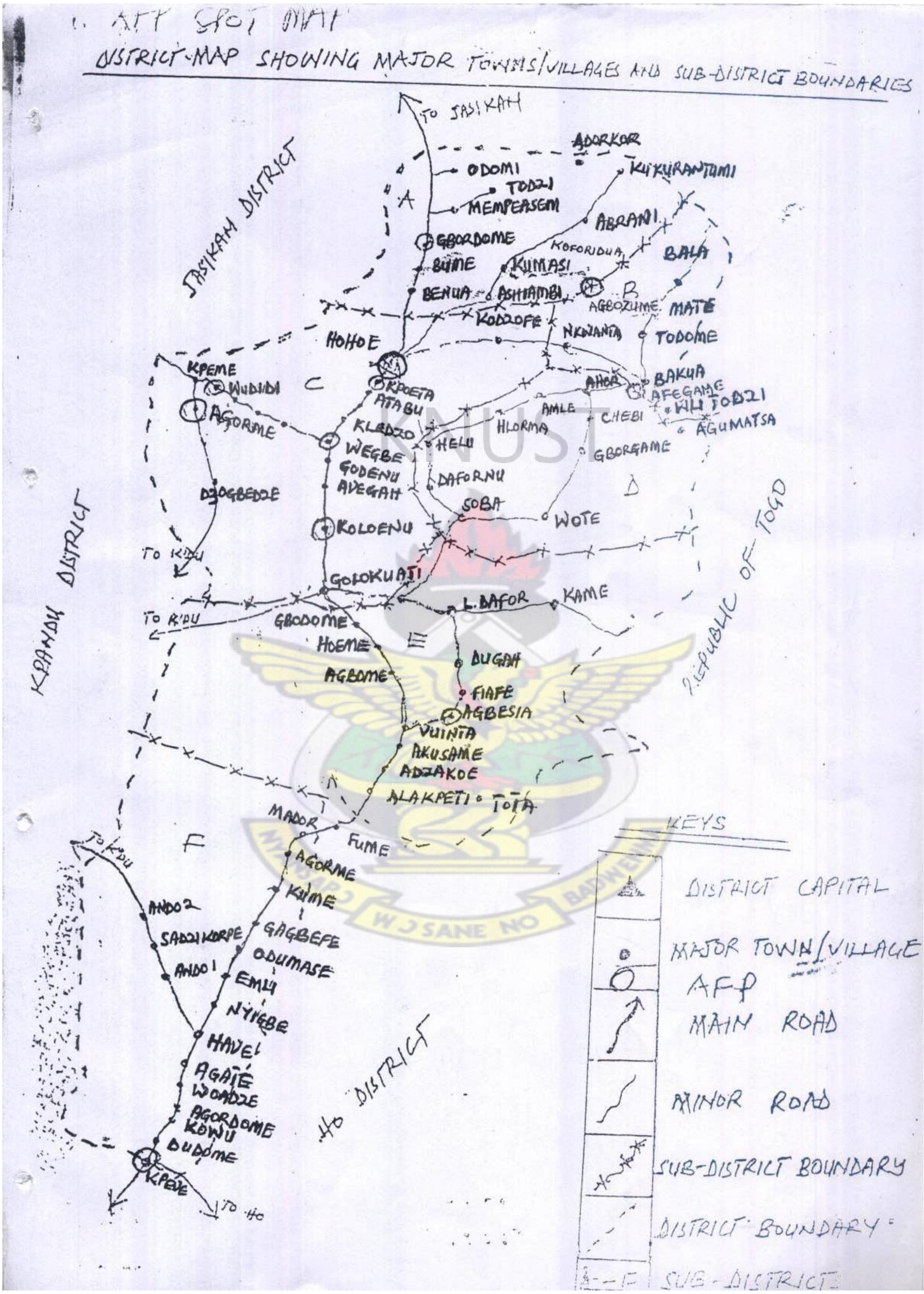
Smith, G.C. (2006). Pregnancy Associated with Plasma Protein A and Alpha-fetoprotein A and prediction of adverse perinatal outcome, Obstet Gynecol 2006; 107(1): 161-166.

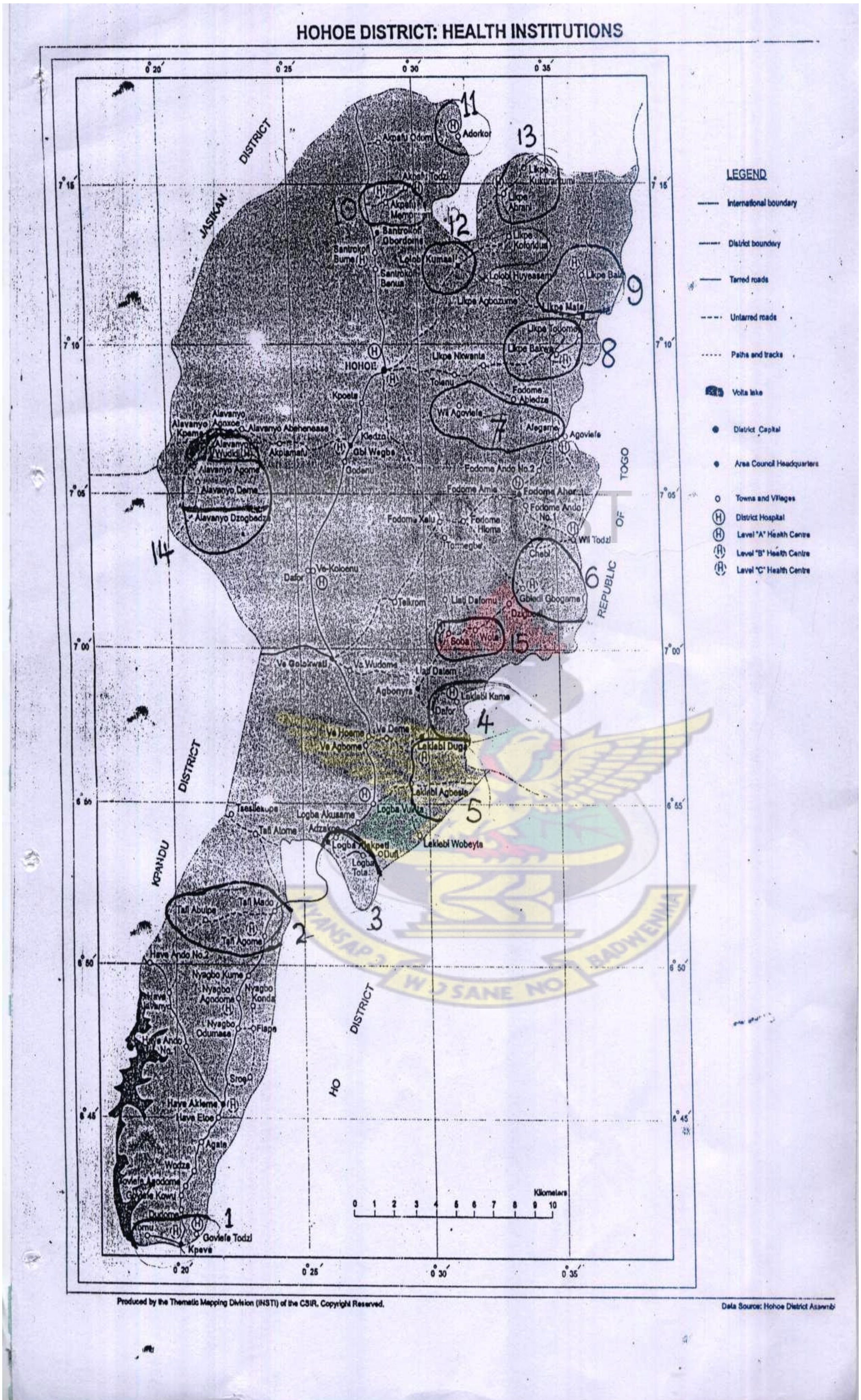
Stacey, T. (2011). Association Between Maternal Sleep Practices and Risk of Late Stillbirth: Article in CME, 17th May 2011.

UMDNJ-Rober, W. J. (2006). Prenatal Tobacco Use and Risk of Stillbirth, Oxford Journals, Vol. 10, Issue 1, pp 159-166.

Yatich, N. J. (2009). Maleria, Intestinal Helminths and other stillbirths in Ghana, Department of Epidemiology, School of Public, University of Alabama at Birmingham(UAB), AL 35294-0022, USA.

APPENDIX A

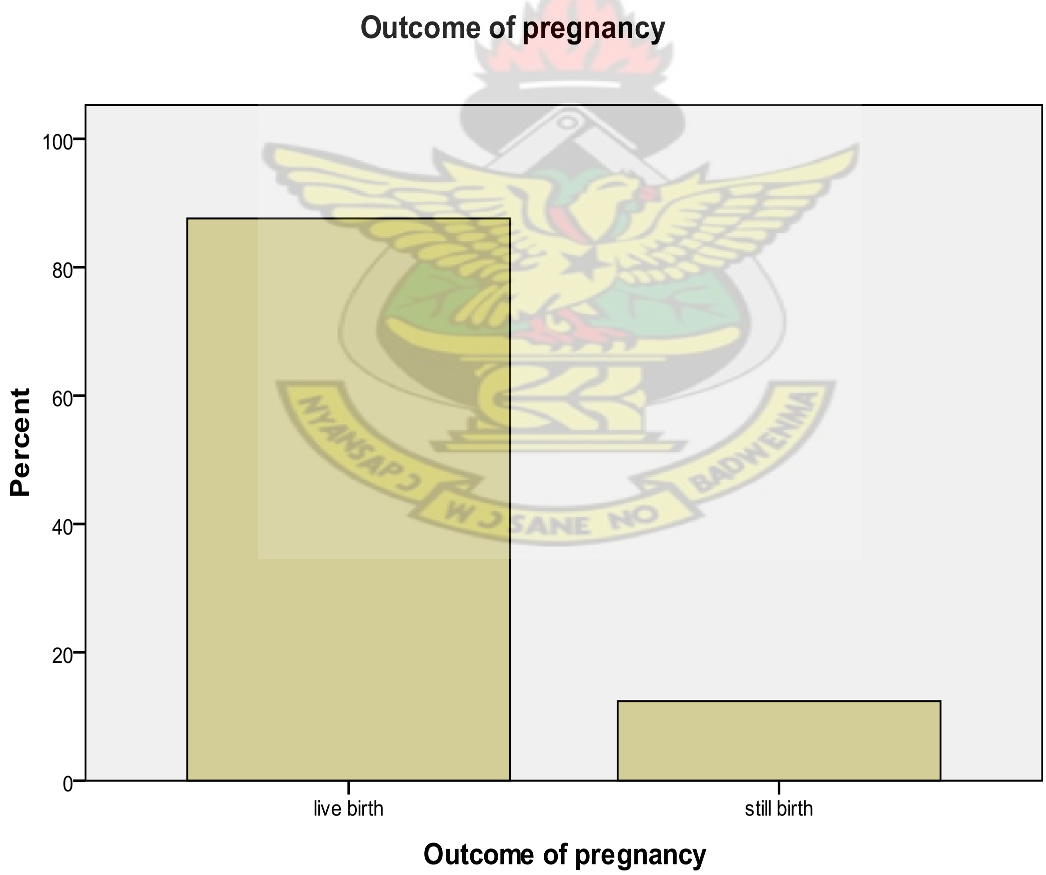

DISTRICT-MAP SHOWING MAJOR TOWNS/VILLAGES AND SUB-DISTRICT BOUNDARIES
TO SASIKAH
SASIKAM DISTRICT
ADORKOR
ODOMI
TODZI
KUKURANTUMI
MEMPEASEM
ABRANI
GBORDOME
KOFORIDUA
BALA
BUME
KUMASI
BENUA
ASHIAMBI
AGBOZUME
MATE
KODZOFE
NKWANTA
TODOME
HOHOE
KPEME
WUDIDI
C
KPOETA
ATABU
AHOR
BAKUA
AFEGAME
WLI TODZI
AGUMATSA
AGORME
KLEDZO
AMLE
HLORMA
CHEBI
HELU
GBORGAME
WEGBE
GODENU
AVEGAH
DAFORNU
DZOGBEDZE
SOBA
WOTE
D
KOLOENU
TO K'DU
KRANDU DISTRICT
GOLOKUATI
REPUBLIC OF TOGO
L. BAFOR
KAME
TO K'DU
GBODOME
HOEME
E
AGBOME
DUGAH
FIAFE
AGBESIA
VUINTA
AKUSAME
ADZAKOE
ALAKPETI
TOTA
MADOR
FUME
F
TO K'DU
AGORME
KUME
ANDO 2
SADZIKORPE
GAGBEFE
ODUMASE
ANDO 1
EMU
NYIGBE
HAVE
AGATE
WOADZE
AGORDOME
KOWU
DUDOME
KPEVE
TO HO
HO DISTRICT
KEYS
DISTRICT CAPITAL
MAJOR TOWN/VILLAGE
AFP
MAIN ROAD
MINOR ROAD
SUB-DISTRICT BOUNDARY
DISTRICT BOUNDARY
A - F SUB - DISTRICT


HOHOE DISTRICT: HEALTH INSTITUTIONS
LEGEND
International boundary
District boundary
Tarred roads
Untarred roads
Paths and tracks
Volta lake
District Capital
Area Council Headquarters
Towns and Villages
District Hospital
Level "A" Health Centre
Level "B" Health Centre
Level "C" Health Centre
JASIKAN DISTRICT
KPANDU DISTRICT
HO DISTRICT
REPUBLIC OF TOGO
HOHOE
Kilometers
0 1 2 3 4 5 6 7 8 9 10
Produced by the Thematic Mapping Division (INSTI) of the CSIR. Copyright Reserved.
Data Source: Hohoe District Assembl
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

APPENDIX B

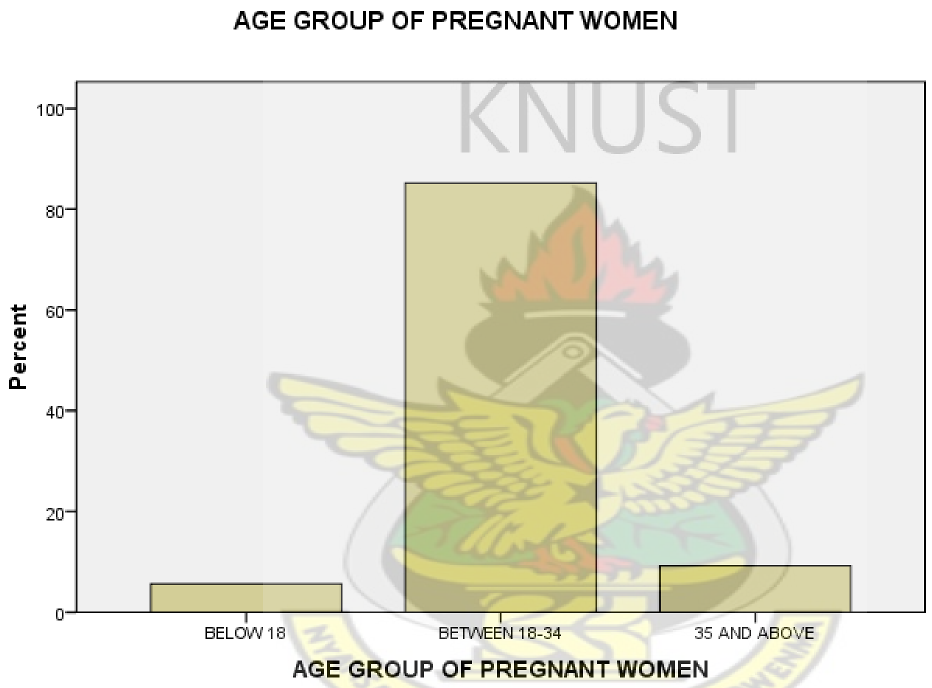
Outcome of pregnancy

		Frequency	Percent	Valid Percent
Valid	live birth	219	87.6	87.6
	still birth	31	12.4	12.4
	Total	250	100.0	100.0



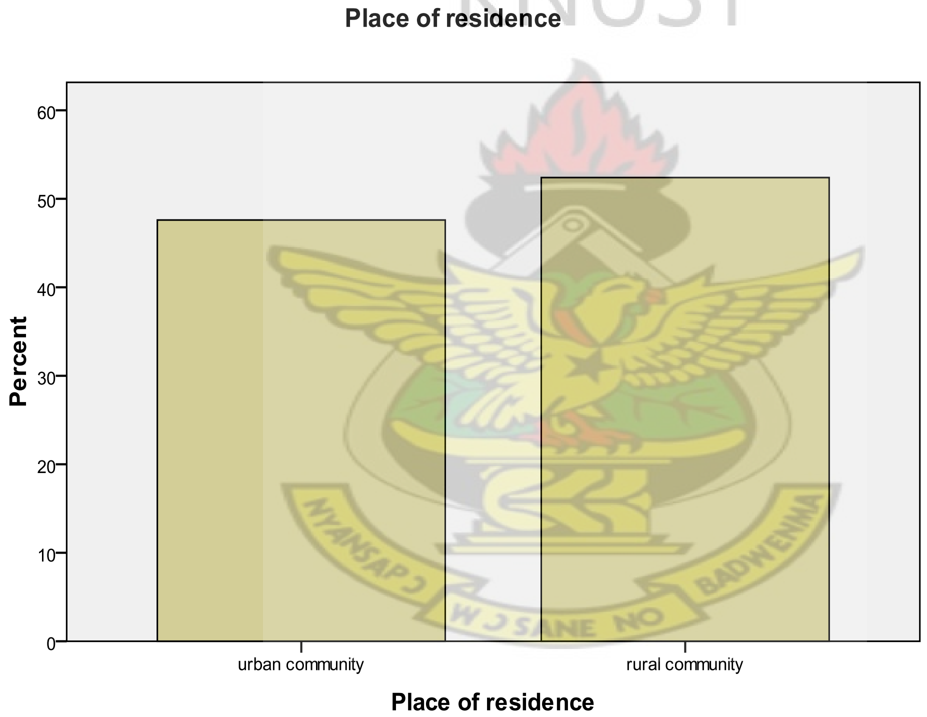
Age group

	Frequency	Percent	Valid Percent
BELOW 18	14	5.6	5.6
BETWEEN 18-34	213	85.2	85.2
35 AND ABOVE	23	9.2	9.2
Total	250	100.0	100.0



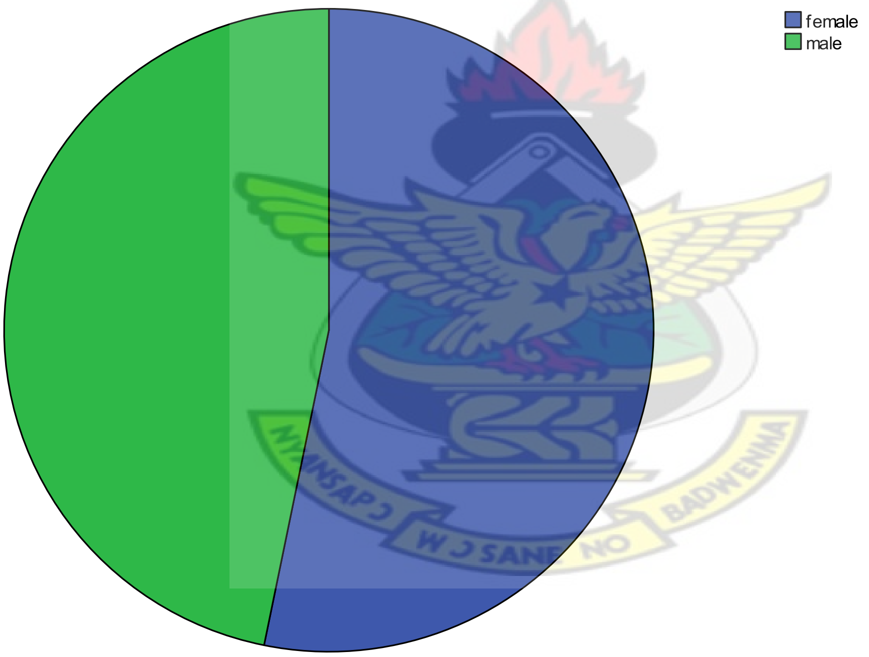
Place of residence

	Frequency	Percent	Valid Percent
urban community	119	47.6	47.6
rural community	131	52.4	52.4
Total	250	100.0	100.0



Sex of baby

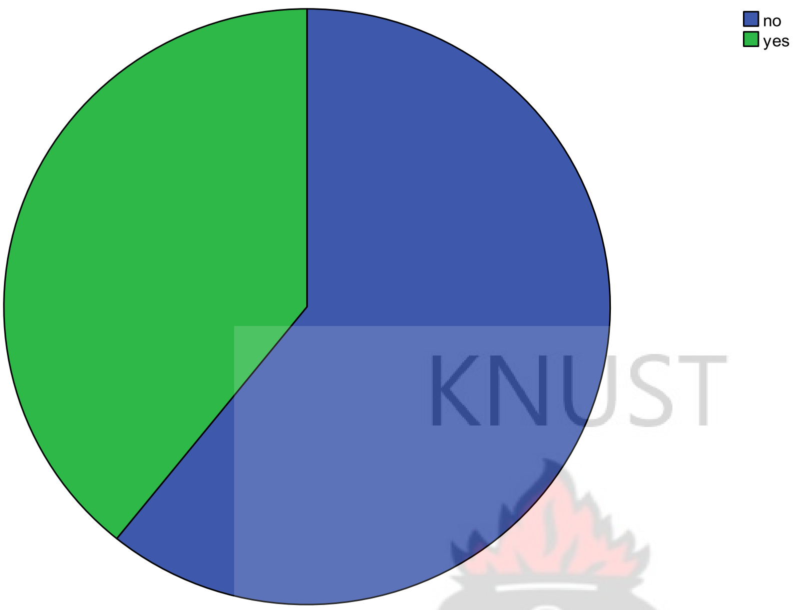
	Frequency	Percent	Valid Percent
female	133	53.2	53.2
male	117	46.8	46.8
Total	250	100.0	100.0



Obstetric problem

	Frequency	Percent	Valid Percent
no	152	60.8	60.8
yes	98	39.2	39.2
Total	250	100.0	100.0

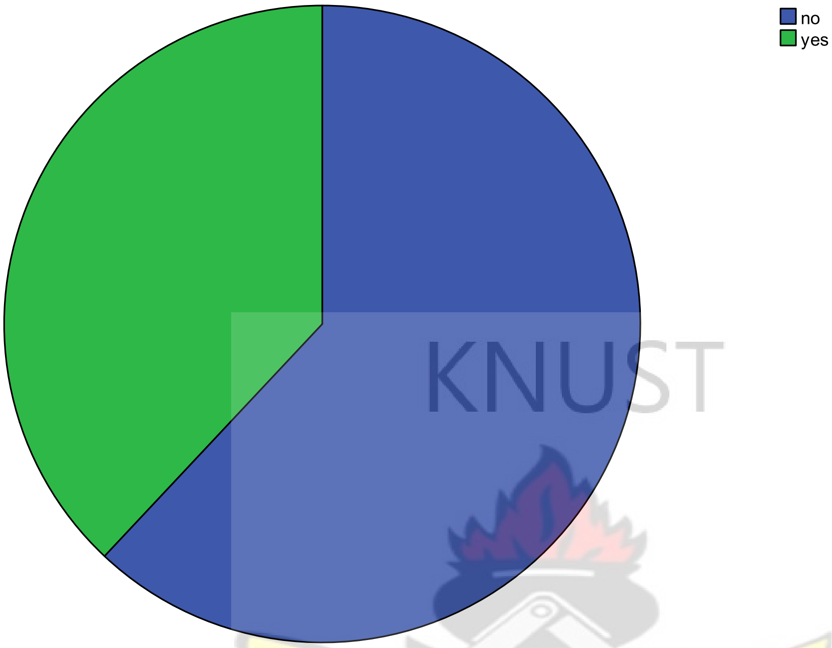
Obstetric problem



Alcoholic drink

	Frequency	Percent	Valid Percent
no	155	62.0	62.0
yes	95	38.0	38.0
Total	250	100.0	100.0

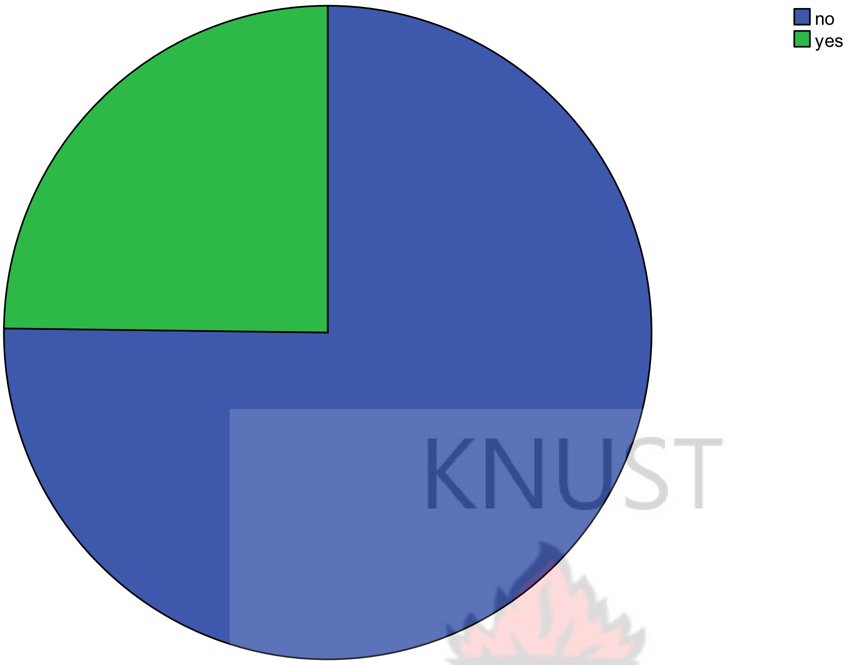
Alcoholic drink



Self-medication

	Frequency	Percent	Valid Percent
no	188	75.2	75.2
yes	62	24.8	24.8
Total	250	100.0	100.0

Self-medication



Sex of baby * Outcome of pregnancy Crosstabulation

			Outcome of pregnancy		
			live birth	still birth	Total
Sex of baby	female	Count	121	12	133
		Expected Count	116.5	16.5	133.0
	male	Count	98	19	117
		Expected Count	102.5	14.5	117.0
Total		Count	219	31	250
		Expected Count	219.0	31.0	250.0

	Frequency	Percent	Valid Percent	Cumulative Percent
6	1	.4	.4	.4
7	14	5.6	5.6	6.0
8	41	16.4	16.4	22.4
9	49	19.6	19.6	42.0
10	51	20.4	20.4	62.4
11	42	16.8	16.8	79.2
12	50	20.0	20.0	99.2
13	1	.4	.4	99.6
14	1	.4	.4	100.0
Total	250	100.0	100.0	

	Frequency	Percent	Valid Percent	Cumulative Percent
1	2	.8	.8	.8
2	25	10.0	10.0	10.8
3	37	14.8	14.8	25.6
4	21	8.4	8.4	34.0
5	25	10.0	10.0	44.0
6	46	18.4	18.4	62.4
7	13	5.2	5.2	67.6
8	55	22.0	22.0	89.6
9	12	4.8	4.8	94.4
10	4	1.6	1.6	96.0
11	2	.8	.8	96.8
12	4	1.6	1.6	98.4
13	2	.8	.8	99.2
24	2	.8	.8	100.0
Total	250	100.0	100.0	

APPENDIX C

DEPARTMENT OF MATHEMATICS

RESEARCHER QUESTIONNAIRE

TOPIC: RISK FACTORS ASSOCIATED WITH STILLBIRTH, A CASE STUDY IN HOHOE MUNICIPALITY, GHANA.

The following questionnaire has been produced as part of a research thesis regarding Factors Affecting Birth Outcomes in Hohoe Municipality of the Volta Region, Ghana. All information received will be treated with strict confidence and individual survey results will not be reported. This questionnaire is restricted to expectant mothers.

SECTION A: Mother's Socio-Demographic Characteristics

1. Age group: [ ] below 18 years [ ] 18 – 34 years [ ] 35 years and above
2. Marital status: [ ] married [ ] single [ ] divorced [ ] widowed [ ] cohabitation
3. Your highest level of education: [ ] no education [ ] primary school [ ] junior high school

   [ ] senior high school [ ] tertiary [ ] other, specify………………..
4. Places of residence: [ ] within an urban community [ ] within a rural community
5. Type of residence: [ ] compound house [ ] self-contain [ ] semi-detached

   [ ] others, specify……………………
6. Source of water: [ ] borehole [ ] well [ ] stream [ ] pipe-borne
7. How many minutes does it take from your home to fetch water from your drinking source?.....................

8. Are you gainfully employed? [ ] yes [ ] no

9. In your estimation, how much do you earn in a month? [ ] GH¢60.00 or less [ ] more than GH¢60.00

10. What is your main occupation?..

11. Is your partner gainfully employed? [ ] yes [ ] no

12. How many children do you already have?...

13. What is your family size?..

14. What type of fuel do you use for cooking? [ ] firewood [ ] charcoal [ ] gas

    [ ] Other, specify…………………………………………………………..

15. Is your partner polygamous? [ ] yes [ ] no [ ] not sure

SECTION B: Mother's Reproductive-Obstetrical information

16. Date of delivery……………………………………………………………….

17. Apgar score…………………………………………………………………

18. First report for antenatal care: [ ] first trimester [ ] second trimester [ ] third trimester

19. Last report for antenatal care before delivery: [ ] first trimester [ ] second trimester

    [ ] third trimester

20. How many times did you visit the hospital for antenatal care?...

21. How old was your pregnancy before delivery?...

22. Weight at first trimester…………………………………………..

23. Weight at last trimester……………………………………………

24. Height of the mother ……………………………………………..

25. Hb: first trimester…………….. second trimester ……………… third trimester………………

26. Weight of baby at birth……………………………………………….

27. Place of delivery: [ ] hospital [ ] delivery home (TBAs) [ ] other, specify……………

28. Sex of baby: [ ] male [ ] female

29. What is the interval between your previous live birth and the current one?…………….

30. Mode of delivery: [ ] caesarean [ ] normal delivery [ ] other

31. Outcome of pregnancy: [ ] stillbirth [ ] live birth [ ] other

32. If live birth, specify type*: [ ] survived after 28 days [ ] dead on/before 28 days (tel. no…………………………..)

33. Did the mother survive? [ ] yes [ ] no

34. Do you have any history of stillbirth? [ ] yes [ ] no

35. How many days did you spend in the hospital/clinic before delivery?.......................

36. How many days did you spend in the hospital/clinic after delivery?..............................

37. During the course of your pregnancy, did you experience any kind of obstetric anomalies? [ ] yes [ ] no

38. If yes, what obstetric problem was it?...

39. What was your general state of health during the course of pregnancy? [ ] sick [ ] healthy

40. If sick, specify what sickness……………………….

41. What treatment was taken?..

42. Did you use treated mosquito net during the pregnancy? [ ] yes [ ] no

SECTION C: Maternal behaviours

43. How many abortions have you had?...

44. Do you smoke cigarettes? [ ] yes [ ] no

45. Do you use any of these other tobacco products? [ ] snuff [ ] chewing tobacco

[ ] cigar [ ] none

46. Does anyone in your household/workplace smoke? [ ] yes [ ] no [ ] not sure

47. Do you drink alcoholic beverages/fermented drinks? [ ] yes [ ] no

48. If yes, how often? [ ] once a day [ ] more than once a day [ ] once a week

[ ] other, specify ……………………………….

49. Did you practice self-medication during pregnancy? [ ] yes [ ] no

50. Do you chew cola nuts? [ ] yes [ ] no

51. How long hours do you work a day?...................................

52. How many meals do you eat in a day? [ ] one [ ] two [ ] three or more meals

53. In your opinion, do you think you eat balanced diets? [ ] yes [ ] no

SECTION D: Societal Norms

54. From your perspective, what traditional/cultural/religious practices or beliefs inhibit life births?

....

....

....

Thank you for your cooperation.