MODELLING OF FACTORS ASSOCIATED WITH PERINATAL MORTALITY RATES IN TANZANIA- *ANALYSIS OF TANZANIA DEMOGRAPHIC HEALTH SURVEY 2015-16 USING BAYESIAN APPROACH*

MTANI NJEGERE MAGESA
W62/12127/2018

A REPORT SUBMITTED TO THE DEPARTMENT OF PUBLIC & GLOBAL HEALTH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN MEDICAL STATISTICS OF THE UNIVERSITY OF NAIROBI

November 2021

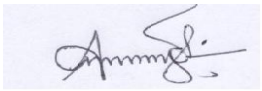
DECLARATION

This report is my original work and has not been submitted for award of a degree in any other University.

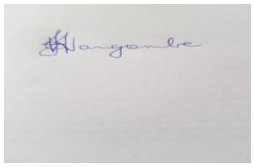
Signature..... Date 11/11/2021

This report has been submitted with our approval as university supervisors.

Henry Kissinger Ochieng Athiany
Lecturer
Department of Statistics and Actuarial Sciences
Jomo Kenyatta University of Agriculture and Technology

Signature:  Date: 11/11/2021

Dr Anne Wang'ombe
Lecturer
School of Mathematics, Chiromo campus
University of Nairobi

Signature:  Date: 11/11/2021

DEDICATION

This project report is dedicated to my family: my wife Joyce Maiga and my children Given, Hellena, Albinus Njegere Jr, David, my mother Hellena Mugeta and my late father Njegere Magesa.

ACKNOWLEDGEMENT

Sincerely from the bottom of my heart, completion of this project would not have been possible without blessings from the Almighty God. Before I go far, I should say Amen! I sincerely would like to pass my gratitude to different people who in one way or another participated or contributed towards the completion of this research. First of all, I would like to pass my special gratitude to my supervisors Dr Henry Kissinger Athiany from Jomo Kenyatta University of Agriculture and Technology (JKUAT) and Dr Anne Wang'ombe (UoN) for their support, guidance, accessibility and valuable critics which contributed towards achieving the present report. They were so cooperative and always calm to me in a way they shaped my proposal and report in the present form**.**

I wish also to express my sincere thanks to my lecturers at the University of Nairobi for the contribution they provided in terms of advice and encouragement. The same also goes to my Manager at Tanzania Medicines and Medical Devices Authority (TMDA) Lake Zone Office, Mrs Sophia Ally Mziray for her motivation, advice and tolerance which helped me to complete my research. Special thanks also go to the management of TMDA for their extended support in terms of finances and material support. I wish also to extend my heartfelt thanks to all my classmates for their cooperation which made my life in Nairobi comfortable.

It is difficult leaving this page without sending profound gratitude to my family especially my wife Joyce Maiga, my children Given, Hellena, Albinus Njegere Jr, David and my mother Hellena for the cooperation they were giving me.

LIST OF ABBREVIATIONS AND ACRONYMS

ANC	Antenatal Care
AOR	Adjusted Odds Ratio
Birthint	Pregnancy interval
BMI	Body Mass Index
DHS	Demographic Health Survey
DIC	Deviance Information Criteria
GLM	Generalized linear model
GLMM	Generalized linear mixed effect model
GMRF	Gaussian Markov random field
HGLMM	Hierarchical Generalized Linear Mixed Effect Model
HIV	Human Immunodeficiency Virus
HSSP	Health Sector Strategic Plan
ICC	Intraclass correlation
INLA	Integrated Nested Laplace Approximation
Mateage	Maternal age at birth
Mateducat	Maternal education
MCMC	Markov Chain Monte Carlo methods
MDG	Millennium Development Goals
MLE	Maximum likelihood Estimation
OR	Odds Ratio
PHC	Primary Health Care
PNM	Perinatal Mortality
SDG	Sustainable Development Goals
SEWI	Social-economic wealth index
TDHS	Tanzania Health and Demographics Survey
U5MR	Under-5 Mortality Rates
UN	United Nations
USA	United States of America
WHO	World Health Organization

Table of contents

LIST OF FIGURES

LIST OF TABLES

DEFINITIONS OF OPERATIONAL TERMS

Perinatal mortality stillbirth and neonatal deaths before 7th birthdate

Bayesian analysis is the inference approach that involves applying prior beliefs to the expected data.

Antenatal care is the routine health control of the presumed pregnant woman.

Intrapartum the period spanning childbirth from the onset of labour to delivery.

Stillbirth loss of a baby before or during delivery

The prior distribution is the probability distribution that expresses one's belief about the quantity before some evidence is taken into account.

Posterior distribution
is the probability that is assigned after the relevant evidence is taken into account.

Maximum likelihood estimation
is a method of estimating the parameters of a probability distribution by maximizing a likelihood function so that the observed data is most probable.

Intra-class correlation
is a measure of how strongly units in the same group resemble each other.

Credible intervals are the probability that the true unobserved value would lie within the interval, given the evidence provided by the observed data.

Neonatal deaths death of a baby before 7 days after delivery

Birth order Order of birth of a child in a sequence of delivery.

ABSTRACT

Background: Perinatal mortality is a big problem facing child survival in Tanzania. The problem has been empirically associated with numerous factors. The challenge which exists is the statistical association of these factors with perinatal mortality. This study aimed to investigate the determinants of perinatal mortality in Tanzania based on the TDHS 2015-16.

Methodology: The study used data from the Tanzania Demographic Health Survey whereby an assessment of the determinants of perinatal mortality was carried out using bivariate chi-square test, normal Bayesian logistic regression and multilevel Bayesian analysis in R-INLA. The multilevel relative risk was computed using log-binomial models in R-INLA. The chi-square test and regression models were undertaken in R version 4.0.3.

Results: A total of 13,266 women aged between 15-49 years and 10,233 children born in the previous 5 years period were included in the study. Overall, 187 stillbirths and 215 neonatal deaths were recorded in the study period, giving a total of 402 perinatal deaths. Perinatal mortality was significantly higher in children of higher birth orders (2+) and maternal body mass index 25+, but lower in children with previous birth intervals of 15+ months. In hierarchical Bayesian analysis, body mass index, previous birth interval and birth order were independently associated with perinatal mortality. Birth order 2-3 (adjusted Odds Ratio (AOR) 5.19 95% CI (5.17,5.21), birth order 4-5 AOR 3.39 95% CI (3.37,3.40), birth order 6+ AOR 5.43 95% CI (5.40,5.46) and body mass index 18.5-24.9 (adjusted Odds Ratio (AOR) 2.22 95% CI (2.20,2.24), body mass index 25-29.9 AOR 2.86 95% CI (2.81,2.91) and body mass index >30 (AOR 2.86 95% CI (2.83,2.89) were associated with increased risk of perinatal mortality. On the other hand, pregnancy intervals 15-26 months, 26-38 months and 39+ months were associated with lower risk of perinatal mortality with AORs

0.336 95% CI (0.335,0.338), 0.340 95% CI (0.338,0.341) and 0.099 95% CI (0.096,0.102) respectively.

Conclusion: Perinatal mortality remains a significant problem in Tanzania. We have identified some of the risk factors, indicating the need to invest in educating mothers about previous childbirth intervals, birth order and body mass index. Reduction of perinatal mortality should also be attributed to interventions that focus on the mother's occupation, age at first birth and education level. Health planners and managers could consider putting a strategic plan in action to assist in improving maternal health in Tanzania, thereby reducing the risk of perinatal deaths.

CHAPTER ONE: INTRODUCTION

1.1 Background

Perinatal mortality (PNM) is defined as the risk for a fetus dying before being born or a neonate dying before the 7th day of his/her birthdate. The perinatal period commences at 22 weeks after gestation and ends 7 completed days after birth (Zupan & Åhman, 2006). It, therefore, refers to the number of stillbirths, and the death of neonates one week after birth. Perinatal mortality rates in any nation are considered to be one of the important indicators of the social-economic welfare (Ebenezer et al., 2019). This is because poor quality care contributes to morbidity and mortality. Over years the perinatal mortality rates in developing countries have been on the rise at an alarming level in contrast to the developed countries (Ebenezer et al., 2019; Racape et al., 2016). Recently, researchers have shown a slowly decreasing trend in perinatal mortality in many sub-Saharan countries over the past 20 years (Ikua A.W, 2010), while the converse is true for the under-5 mortality within the same region (Ikua A.W, 2010).

Globally, perinatal mortality is a public health concern, although it is most pronounced in developing countries (Zupan & Åhman, 2006). In 2009, globally there were 2.6 million stillbirths recorded, and in 133 million live births in the same year, 2.8 million died in the first week of life (Frøen et al., 2016). This translated to 7,120 stillbirths and 7,670 neonatal deaths per day as well as 40.6 deaths per 1000 live births. It is also estimated that about 66% of the stillbirths globally occur only in 10 countries namely India, Pakistan, China, Nigeria, Bangladesh, Ethiopia, Democratic Republic of Congo, Indonesia, Afghanistan and Tanzania (Frøen et al., 2016). The patterns of these deaths are similar to under 5 mortalities, with the majority occurring in the developing countries(Ogbo et al., 2019).

Since Tanzania is one of the 10 most-affected countries globally, the global perinatal mortality trend is similar to the situation in Tanzania. The Tanzania Demographic Health Survey 2015-2016 has reported perinatal mortality of 39 deaths per 1000 live births(Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) & ICF, 2016). The government of Tanzania in collaboration with other partners has conducted several interventions to reduce perinatal mortality in the country among other health-related problems (Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC), 2015). However, due to lack of funding, the perinatal mortality rates continue to be high in the country (Ogbo et al., 2019). For example, under-5 mortality in Tanzania in the period between 2004 and 2016 was almost halved while perinatal mortality remained high (Ogbo et al., 2019). The World Health Organization (WHO) estimated and placed Tanzania in a group of countries with perinatal mortality rates between 40-59 by the year 2000 (Zupan & Åhman, 2006). The hit of 39 deaths per 1000 live births in 2015-2016(Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) & ICF, 2016) indicates slow and little progress.

The millennium development goals (MDGs) for poverty reduction which were adopted by world countries in 2000 place public health improvement as a central focus(Jeffrey D. Sachs, 2003; United Nations (UN), 2015). This is because public health plays a pivotal role in the poverty reduction (Jeffrey D. Sachs, 2003). In the MDG-4, reduction of infant and child mortality rate

by two-thirds by 2015 in the developing countries was emphasized (Jeffrey D. Sachs, 2003). Although this was not achieved, there was a notable decline in under-5 mortality from 90 deaths per 1000 in 1990 to 43 deaths per 1000 in 2015 (Nations Unies, 2015). In an attempt to meet MDG-4, Tanzania as a member of the United Nations (UN) has been implementing different programs to improve health in the country including building health centres for every ward in line with the implementation of the Health Sector Strategic plan 2015-2020 (HSSP 1V) (Ministry of Health, Community Development, Gender, Elderly and Children, 2015). Despite all these efforts, Tanzania has achieved slow gain in reduction of Perinatal mortality (Mbaruku et al., 2009).

Of particular concern is the absence of stillbirths in millennium development goals and the Sustainable Development Goals, making stillbirth a neglected issue, invisible in policies and programs (Frøen at al., 2016). This fact should attract the governments in developing countries to set priorities and resources towards combating stillbirth to improve life.

Thus, identifying the risk factors associated with perinatal mortality is very important in designing programs to combat the problem. The fact of life is that perinatal deaths will continue to occur and can be reduced if the risk factors are given high priority in antenatal care and intrapartum programs (Zupan & Åhman, 2006). Therefore, this study is aiming at modelling different variables for perinatal mortality in Tanzania using the Tanzania demographic and health survey data of 2015-2016.

1.2 Statement of the Problem

Perinatal mortality is a global issue that primarily affects developing countries, especially the Sub-Saharan region. The global trend shows that there has been a slow decline in perinatal mortality in many developing countries(Zupan & Åhman, 2006). The figure below shows regional perinatal mortality trends. From *Figure 1*, it appears that Sub-Saharan

countries have higher perinatal mortality rates than the rest of the world. These statistics show the heavy load of childhood mortality created by perinatal mortality, especially in the Sub-Saharan region.

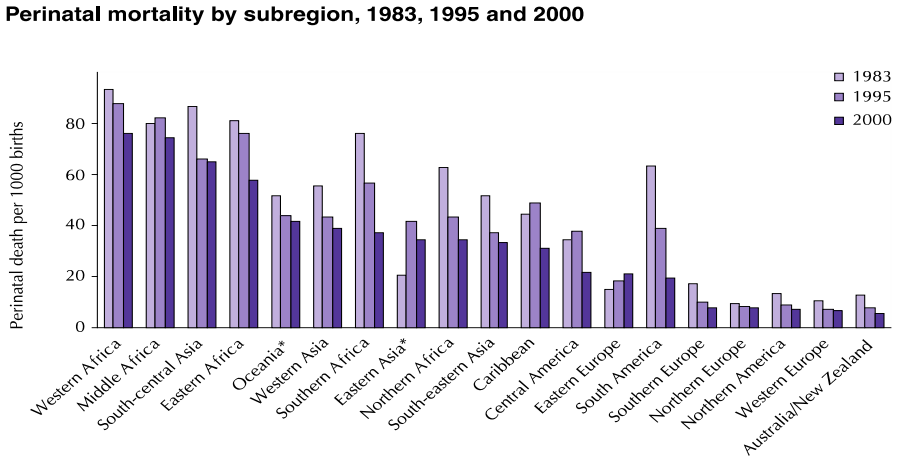


Figure 1: Global perinatal mortality by region

Source: WHO (2006).

Moreover, previous literature has been focusing on determinants of infant and child mortality; and very few have concentrated on perinatal mortality which significantly contributes to infant and child mortality (Akombi & Renzaho, 2019). In addition, most of these studies have used hospital or community-based data (Gabrysch et al., 2019; Kidanto et al., 2006; Mbaruku et al., 2009; McDermott et al., 1996) and very limited studies have used demographic health survey data for analysis (Biracyaza E* and Habimana S, 2019; Chinomona & Mwambi, 2015). The results from the hospital and community-based studies have limitations in that they cannot be used to infer the general population. Previous research has also shown that there are many unreported cases of perinatal mortality (Anwar et al., 2018; Cardoso et al., 2016). This poses a challenge in the analysis of studies that would be used by policy-makers and implementers to set public health interventions aiming at reducing perinatal mortality. Thus, this study can go a long way in advising the policy-makers on the priority areas where the

limited resources should be directed. Lastly, many studies which involve hospital and community-based data and even demographic surveys have used logistic regression analysis to determine the effect of selected variables on dichotomous outcomes. However, this only makes it possible to estimate the intercept and other regression coefficients but leaves room for uncertainty in the estimates. This is because there could be several reasons which can lead to uncertainty with the empirical data such as unreported cases of stillbirth and neonatal deaths and regional differences of perinatal mortality within the country. By employing normal logistic regression, tends to ignore these prior beliefs about the data. Hence, a Bayesian approach will help in improving the estimates by incorporating prior knowledge in the model to obtain the estimates.

1.3 Justification of the study

Perinatal mortality is a major global problem contributing to a significant number of stillbirths and neonatal mortality. One of the greatest challenges is that there has been a gradual decline in perinatal mortality in most Sub-Saharan countries including Tanzania. Most studies in the field of child mortality have only focused on under-5 mortality and researchers have not treated stillbirth in much detail. On the other hand, many types of research in hospital and community-based and few demographic survey studies have been restricted to classical logistic regression analysis. To date, only a limited number of studies have been identified to use Bayesian logistic regression analysis. Therefore, modelling of risk factors associated with perinatal mortality in Tanzania using the Bayesian multivariate logistic regression model will help in identifying the risk factors which significantly contribute to perinatal mortality by incorporating prior knowledge in the model.

1.4 Research questions

(i) What is the prevalence of perinatal mortality in Tanzania among the selected variables?

(ii) What are the variables associated with perinatal mortality in Tanzania?

(iii) What is the relative risk of perinatal mortality in Tanzania?

1.5 Objectives

1.5.1 General Objective

The main objective of this study was to determine the factors associated with perinatal mortality in Tanzania based on Tanzania Demographic and Health Survey 2015-2016 data.

1.5.2 Specific Objectives

(i) To determine the prevalence of perinatal mortality in Tanzania

(ii) To determine the bivariate association between maternal education, place of residence, maternal age at birth, maternal age at first birth, birth interval, marital status, maternal BMI, maternal weight, birth order, child's sex, maternal occupation and social-economic status and the risk of perinatal mortality in Tanzania using a Bayesian approach.

(iii) To estimate the relative risks of perinatal mortality in Tanzania using the selected socio-economic and socio-demographic factors of the population, accounting for potentially complicated dependency structures of the population using a suitable hierarchical Bayesian Approach.

CHAPTER TWO: LITERATURE REVIEW

2.1 Perinatal mortality

Perinatal mortality is one of the key indicators of the development of any country (Ebenezer et al., 2019; Zupan & Åhman, 2006). A high level of perinatal mortality indicates a low level of development of a community, and also low perinatal mortality signifies improved health systems of the community in question (Zupan & Åhman, 2006). Perinatal mortality rates may be attributed to several factors such as child sex, birth order of a child, maternal education etc.(Biracyaza E* and Habimana S, 2019). Implementation of Primary Health Care (PHC) for all by the year 2000 (WHO, 1978) aimed at reducing neonatal mortality rates globally including Tanzania. However, stillbirth has continued to remain high in Tanzania (Kidanto et al., 2006).

Child mortality has been a very critical issue in evaluating the progress or achievement towards the Millennium Development goal 4 (MDG-4) which required member countries to reduce child mortality by two-thirds between the period 1990 to 2015(Mejía-Guevara et al., 2019). The Sustainable Development Goal 3 (SDG 3) calls for member countries to reduce infant mortality to 12 deaths per 1000 live births and under-5 mortality to as low as 25 deaths per 1000 live births(United Nations, 2015). Even though MDG-4 and SDG-3 do not put any emphasis on the reduction of stillbirth, individual countries should focus on the reduction of stillbirth and neonatal mortality. The efforts towards achieving this goal require accurate statistics for mortality measurements (Rerimoi et al., 2019). Additionally, while the target for MDG-4 was a reduction of child mortality to two-thirds, this was not achieved by many member states resulting in a reduction by only a half of the global child mortality (United Nations (UN), 2015).

According to the Tanzania Ministry of Health, there has been a remarkable decline in infant and child mortality rates in Tanzania (Ministry of Health,

Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) & ICF, 2016). The UN report estimates that only Tanzania, Kenya, Rwanda, Uganda and Senegal can achieve a reduction of under-5 mortality rates (U5MR) to 25 deaths per 1000 live births by 2030(Mejía-Guevara et al., 2019). The report goes further and suggests that only Tanzania and Rwanda can achieve a reduction of neonatal mortality rate to 12 deaths per 1000 live births by the year 2030(Mejía-Guevara et al., 2019).

2.2 Causes of Perinatal Mortality

There are several causes of perinatal mortality which can be categorized into three major parts: Those which happen in the period before childbirth (antepartum), those which occur during childbirth (intrapartum) and those related to neonatal mortality. Among other many causes of perinatal mortality, antepartum and intrapartum complications are the major causes of stillbirth (Afshan et al., 2019). Other causes include a congenital abnormality, intrapartum related asphyxia and high-risk fertility behaviour (Becher et al., 2004).

Diseases like HIV and malaria have also been related to poor pregnancy outcomes (Akombi & Renzaho, 2019; McDermott et al., 1996). The interaction of HIV and other maternal infections may result in poor birth outcomes. For example, Lesotho is among the countries with the highest perinatal mortality in Sub-Saharan Africa with also a high level of HIV infection (Akombi & Renzaho, 2019). The high mortality rates in Lesotho are thought to be attributed to HIV infection. Studies have shown that maternal HIV status leads to an increased risk of perinatal mortality (Akombi & Renzaho, 2019; Chinomona & Mwambi, 2015). Also, HIV

leads to the low birth weight which may increase the risk of perinatal mortality, especially during the neonatal age (Akombi & Renzaho, 2019).

2.3 Determinants of Perinatal Mortality

Generally, studies have identified several determinants of perinatal mortality such as type of residence, maternal education, birth interval, child sex, mother's weight and child-size (Conde-Agudelo et al., 2005; Gabrysch et al., 2019; Ikua A.W, 2010; Jacobsson et al., 2004; Kidanto et al., 2006; Muttarak & Dimitrova, 2019; Racape et al., 2016). All these fall under three categories which are social-economic factors, community factors and proximate factors. Type of residence covers several attributes which have a direct effect on the well-being of individuals. It has a connection to weather, hygiene, proximity to health services, infrastructures, altitude, water sanitation and social-economic status. For example, areas that are affected by extreme weather conditions like drought, floods amongst others have recorded high child mortality rates since it may negatively affect nutrition in the area (Muttarak & Dimitrova, 2019).

Minsart et al., 2013, showed in their work that perinatal mortality varied according to the origin of the mother and her naturalization status. In this study, it was reported that naturalized immigrants who had access to good healthcare had low perinatal mortality than those who were not naturalized. People living in rural areas and other disadvantaged areas are likely to experience high neonatal mortality rates and stillbirth due to lack of maternal-child health services (de Graaf et al., 2013). In their study, they reported a high risk of adverse perinatal mortality (21%) in deprived districts.

In addition, there have been always rural-urban differences in terms of accessibility, readiness and delivery of health services among many countries in the Sub-Saharan Africa (Kanyangarara et al., 2018). A recent

study by Lisonkova et al., 2016 has clearly shown how the place of residence impacts perinatal mortality through the following results: severe maternal morbidity and rural residence (AOR 1.15,95% CI 1.03-1.28) compared to urban women. This study also indicated that rural women had significantly higher rates of eclampsia (AOR 2.7,95% CI 1.79-4.08), obstetric embolism (AOR 2.16, 95% CI 1.14-4.07), uterine rupture (AOR 1.96, 95% CI 1.42-2.72). The same study also reported high neonatal morbidities in infants born in rural areas compared to those born in urban areas.

Maternal level of education is also a determinant of neonatal mortality and stillbirth. This has been shown by numerous studies indicating that mothers with low education have a high probability of their baby dying in childhood than the highly educated one (Auger et al., 2012; Kiross et al., 2019a; Rahman et al., 2010; Wehby & López-Camelo, 2017). This implies that the level of education of women of childbearing age is associated with child survival or mortality and the likelihood of stillbirth. The general concept is for the policy-makers to invest in maternal education to reduce child mortality and stillbirth (Kiross et al., 2019a)(Kiross et al., 2019a). One notable study was conducted in Qatar by Rahman et al., 2010 where they showed how improved maternal education has reduced neonatal mortality rates from 26.27/1000 in 1974 to 4.4/1000 in 2008, perinatal mortality from 44.4/1000 in 1974 to 10.58/1000 in 2008 and maternal mortality rate to zero.

Birth interval has also been identified by several studies to be one risk factor for stillbirth and infant-child mortality in many countries (Becher et al, 2004; Molitoris et al., 2019; Mondal et al., 2009). This means that if children are not spaced enough, there is a high risk for death of a child and stillbirth due to maternal factors (Molitoris et al., 2019). Andargie et al., 2013 clearly described the effect of short birth intervals on perinatal

mortality whereby the AOR for short inter-pregnancy less than 24 months was 2.58 compared to more than 24 months interval in a study conducted in Ethiopia. This outcome is supported by another study conducted in Latin America (Conde-Agudelo et al., 2005 which reported infants with short inter-pregnancy intervals between 18-23 months to have morbidities like early neonatal death, fetal death, low birth weight, very low birth weight, preterm death, very pre-term birth and small for gestational age.

Mother's age at first birth also has a very big impact on perinatal mortality. Several types of research have nailed on the importance of reducing adolescent pregnancy to curb the problem of neonatal deaths(Jacobsson et al., 2004; Neal et al., 2018; United Nations, 2015). Neal et al., 2018 reported an increased risk of neonatal mortality among infants born from mothers under 16 years old. Although this study reported the risk of neonatal mortality among adolescents, generally neonatal mortality has been attributed to low age and advanced age at birth. Studies have shown that infants born to mothers above 40 years old have a much higher risk of neonatal mortality than those born to mothers less than 40 years old (Jacobsson et al., 2004; Mutz-Dehbalaie et al., 2014). This concludes that perinatal mortality, intrauterine fetal death and neonatal deaths increase with adolescence age and advanced age.

Social-economic status plays a pivotal role in determining stillbirth and infant mortality in low-middle income countries. Access to health services, health infrastructures, maternal education, maternal nutrition, infant nutrition amongst others is well influenced by the economic wealth of the community or household. Studies have shown that stunted under-5 children are mainly from poor families than wealthy families. For example, stunting was observed to be 45.1% in poor families while it was 26.9% in wealthy families in a study conducted in Ethiopia (Mohammed et al., 2019). In Uganda, the stunting of under-5 years children was 30% and most of the

children were from families whose mothers were widowed, divorced or separated (Ruth Sharon Apio et al., 2019). It has been noted that even utilization of antenatal care services differs in the community based on the social-economic status of the households. For instance, utilization of antenatal care (ANC) services is concentrated in women from wealthier households than in women from poor households (Novignon et al., 2019).

2.5 Bayesian Regression

2.5.1 General Description

The Bayesian regression analysis has gained wide application in recent years due to correct estimates of regression coefficients it produces. Its application is not only in the medical field but it crosses to other fields like spatial statistics, politics, finance, business amongst others. The introduction of prior information without losing posterior distributions gives a useful advantage of producing correct estimates (Pateras, K & Ntzoufras,I, 2013).

Incorporating prior information in a model renders the research more attractive than empirical data through the incorporation of experience, expert opinion or probabilistic thinking of a researcher. This study incorporates some non-informative priors which play a minimal role in the posterior distribution and with the interest of letting the data speak for themselves (Chinomona & Mwambi, 2015; Joanne L. Shin, 2015). Also, the prior knowledge about perinatal mortality will be incorporated in the model using the relevant prior.

2.5.2 Bayes Formula

The general idea in Bayes theorem is:

$$p(Y_i|X_i) = \frac{p(X_i \mid Y_i)p(Y_i)}{p(X_i)}$$

where:

$p(Y_i \mid X_i)$ is the posterior probability of subject i

$p(X_i \mid Y_i)$ is the likelihood

$p(Y_i)$ is the prior probability

$p(X_i)$ is the normalizing constant so that the probability is appropriately scaled within 0 to 1

The term $p(X_i)$ in the equation above is always difficult to compute. Therefore,

$$p(Y_i|X_i) \propto p(X_i|Y_i)p(Y_i)$$

which means the posterior distribution is the product of the likelihood observed from the collected data and the prior distribution.

2.6 Bayesian Logistic Regression Model

In modelling the risk factors associated with perinatal mortality, we are interested in showing the relationship between the risk factors in this case regarded as independent variables and the outcome variable (dependent variable). In the TDHS 2015-2016 data, the outcome variable is perinatal mortality. This implies that the outcome variable is categorical binary (dichotomous) with only two probabilities and the interest is to study how a set of predictor variables are related to dichotomous outcomes(Harrell, 2001). The predictor variables may be either discrete or continuous.

Since the outcome variable is binary, a binary logistic model is appropriate in modelling the risk factors associated with perinatal mortality. The logistic regression model, therefore, provides a method of modelling binary outcomes which takes values 0 or 1. The proportion of children dying is an estimate of probabilities of death in each category. However, this

relationship is always nonlinear and the probabilities of death change slightly at the lower and upper extremities.

The proportion of children dying can be represented by logistic regression represented by a linear function of covariates as follows:

$$\mathrm{p} = \frac{\boldsymbol{e} - \left(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} \ldots\ldots \beta_p x_{ip}\right)}{\mathbf{1} + \boldsymbol{e}^{-\left(\boldsymbol{\beta_0} + \beta_1 \boldsymbol{x_{i1}} + \beta_2 \boldsymbol{x_{i2}} \ldots\ldots \beta_p \boldsymbol{x_{ip}}\right)}}$$

Where β is a vector of regression coefficients.The logistic or logit function is used to transform a nonlinear curve into a linear curve and change a range of proportion from 0 to 1 to -∞ to+∞. The linear function of covariates for the binary outcome will be as follows.

$$logit(Y_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots\ldots \beta_p X_p + \mathcal{E}$$

$$\text{Where } \mathcal{E} \sim \mathrm{N}(0, \sigma^2)$$

and this model was used for normal Bayesian analysis with the application of priors.

Bayesian logistic regression has been used in several research studies to solve different problems. For example, Bayesian logistic regression models were used by Chinomona and Mwambi(Chinomona & Mwambi, 2015) to model the risk of HIV in Zimbabwe using the Zimbabwe DHS dataset 2010-11. Two other studies conducted in Nigeria used the Bayesian logistic regression model to determine Exclusive breastfeeding practice and the risk of malaria (Adigun et al., 2015; Gayawan et al., 2014). Gayawan et al (Gayawan et al., 2014) conducted a stepwise Bayesian analysis to determine exclusive breastfeeding practice in Nigeria whereby the study could differentiate the likelihood of exclusive breastfeeding among the Nigerian states. In a similar context, Adigun (Adigun et al., 2015) used Bayesian geostatistical modelling of the 2010 malaria indicator survey to determine malaria risk in Nigeria.

2.6.1 Likelihood of the Model

For binary outcome studies as in the case of perinatal mortality, we should consider the outcome variable as Bernoulli random variable. If we want to choose values for parameters of logistic regression, we need to use maximum likelihood estimation (MLE). Since we are considering the outcome variable as Bernoulli random variable, then

$$Y_i \sim Bern(p)$$

where $p = \sigma(\theta^T x)$.

Now if the probability of one data point is computed as a Bernoulli probability mass function, the following equation is obtained:

$$P(Y|X) = \sigma(\theta^T x)^y . [1 - \sigma(\theta^T x)]^{1-y}$$

The likelihood for Bernoulli distribution will be given by:

$$L(\theta) = \prod_{i=1}^{n} \sigma(\theta^T x^i)^{yi} . [1 - \sigma(\theta^T x^i)^{(1-yi)}$$

If the logarithm of likelihood $L(\theta)$ is taken, the log-likelihood is as follows:

$$LL(\theta) = \sum_{i=1}^{n} y^{(i)} log\sigma(\theta^T x^i) + (1 - yi)\log[1 - \sigma(\theta^T x^i)]$$

From the likelihood of the model, the value of θ is chosen which can maximize the log-likelihood.

2.6.2 Prior Distribution

2.6.2.1 Informative Priors

With informative priors, the useful domain knowledge about the parameters is encoded. The domain knowledge may arise from experience, from prior knowledge about the parameter or the previous literature. Informative priors

also can be generated from pilot studies as was emphasized by a study by Morris (Morris et al., 2013).

2.6.2.2 Non-Informative Priors

Non-informative priors are used when nothing is known about the value of the parameter. These kinds of priors have minimal effect on the posterior distributions and provide a greater chance for the data to speak for themselves (Chinomona & Mwambi, 2015**).** Since the prior variance is a decreasing function, this means the larger the variance is, the less influential the prior is. A large prior variance normally means a relatively weak prior distribution (M. Zhu & Lu, 2004). Therefore, non-informative priors are characterized with large variance.

2.6.2.3 Conjugate Priors

Conjugate priors are priors with prior distribution and posterior distribution taking the same probability distribution. For instance, if the prior probability $p(\theta)$ is chosen so that the posterior distribution $p(\theta|y)$ is of the same probability distribution, it is called conjugate priors. Beta distribution is a usual conjugate before taking a form Beta (α, β) and in this case α and β are the hyperparameters.

2.7 Choice of priors

To make correct predictions on unseen data, use of prior distributions on the regression parameters which assigns a high probability that most entries will have values at or near zero (Madigan, 2005; Souza & Migon, 2004) in Gaussian surface is encouraged. Since the outcome variable follows the Bernoulli distribution, conjugate priors of the beta distribution are

considered to obtain posterior distribution which takes the form of Beta (α, β).

2.8 Posterior distribution

From the Bayes formula, the posterior distribution is computed from the product of the prior distribution and the likelihood given as:

$$\text{Posterior} = p(X_i|Y_i)p(X_i).$$

Since the likelihood is already established and if the prior is assumed to be beta distribution, then the posterior distribution is obtained by combining the full likelihood and the prior distribution which can be derived as follows:

Posterior for Bernoulli-Beta,

Assuming the posterior of $\theta|x_{1:n}$

$$p(\theta|x_{1:n})$$

$$\propto p(|x_{1:n}|\theta)p(\theta)$$

$$\_\theta \sum x_i \, (1-\theta)^{n-\sum x_i} \frac{1}{B(a,b)} \theta^{a-1}(1-\theta)^{b-1}(0 < \theta <$$

$$\propto \theta^{\varepsilon} + \sum x_i - 1 \, (1-\theta)^{b+n} - \sum x_i - 1/0 < \theta < 1)$$

$$\propto Beta(\theta|a) + \sum x_i \, , b + n - \sum x_i$$

The posterior distribution above from full likelihood and beta prior distribution is not analytically easy to deal with, hence the use of the Markov Chain Monte Carlo (MCMC) algorithm to sample from the posterior distribution (Chinomona & Mwambi, 2015). Calculation of posterior distribution is the main objective of the Bayesian paradigm which is then used to make an inference. Since the obtained distribution is probabilistic, then the derived distribution is called predictive distribution (Gelman et al., 2013).

2.9 Parameter Estimation

Parameter estimation in the Bayesian paradigm involves updating the prior beliefs based on the newly introduced evidence. Having prior knowledge about perinatal mortality in Tanzania helps to extend knowledge to the gathered data to obtain posterior information. One common computation method which is used to obtain the posterior distribution is the Markov Chain Monte Carlo algorithm.

2.9.1 Markov Chain Monte Carlo Algorithm

The Monte Carlo algorithm is simulation-based and uses random numbers generated from some probability density function to generate a posterior distribution (Chinomona & Mwambi, 2015). The method is used to model the probability of different outcomes which cannot be easily predicted due to the intervention of random variables (Gelman et al., 2013; Pateras, K & Ntzoufras,I, 2013; Wakefield, 2013). There are two types of Monte Carlo methods which are Gibb's sampling and Metropolis-Hastings (Pateras, K & Ntzoufras,I, 2013).

2.9.1.1 Gibbs Sampling

Gibb's sampling is a very useful simulation method used to sample from distributions that are difficult to simulate directly. It is a randomized algorithm used to obtain a sequence of observations from a specified distribution. It can be illustrated by assuming the sample is simulated from the posterior distribution of $p(y|x)$ where y is a vector of two parameters y_1 and y_2. To begin the initial values are set for y as ${y^0}_1$ and ${y^0}_2$ and determine posterior distribution through generating ${y_1}^n$ and ${y_2}^n$ where $n = 1,2,3 \ldots \ldots$

Gibb's sampling produces a sequence of $y^n = ({y_1}^n, {y_2}^n) = ({y_1}^n, {y_2}^n)^T$, $n = 1,2 \ldots$ in which each pair is independent of each other and create a chain. The chain of a sequence of independent posterior distribution from

the given number of iterations can produce a summary measure such as the mode, mean and median.

2.9.1.2 Metropolis-Hastings

Metropolis-Hastings is a Markov chain Monte Carlo method for obtaining a sequence of random samples from a probability distribution. The method was introduced by Nicolas Metropolis in 1953 (Chinomona & Mwambi, 2015) and modified further by Hastings in 1970 (Pateras, K & Ntzoufras,I, 2013). The Metropolis-Hastings algorithm begins with an arbitrary choice of initial values of model parameter $\theta_0 = ({\theta_1}^0, \ldots . {\theta_m}^0)$ and this initial value is automatically accepted in the model (Bonamente, 2017). It takes some time and several iterations for the chain to forget its initial position and start sampling posterior distribution.

If we want to determine posterior distribution $p(y|x)$ and that Markov chain is at y^{th} iteration and the next position denoted as $\dot{y}$ will be accepted if it is in the region of high posterior mass or else it is accepted with certain probability (Chinomona & Mwambi, 2015). The density for $\dot{y}$ at n iteration is denoted as $p(\dot{y}|y^n)$. If $\dot{y}$ is accepted, meaning $\dot{y} = y^{n+1}$ then the next move is executed and if rejected the process stops at y^n. The requirement for the move to occur requires that the probability $r = \frac{p(\dot{y}|x)}{p(y^n|x)} > 1$ and if r<1 the move stays at y^n. The move continues until convergence is reached.

2.9.2 Convergence test

The convergence test and stopping time of the chain are critical issues for Monte Carlo Markov Chain. The most common convergence tests which can be used with relative ease are the Geweke z-score test, Gelman-Rubin test and The Raftery-Lewis test (Bonamente, 2017). Others include the trace-plots test and Heidelberg-Welch test (Chinomona & Mwambi, 2015).

Before inference is made to the posterior distribution, one must test if convergence for stationary distribution for MCMC has been reached. Convergence indicates that the chain has started to sample posterior distribution and that the MCMC samples are now representative of the distribution of interest. In addition, convergence justifies that the MCMC samples are not biased by the choice of the initial point of the chain (Bonamente, 2017).

In the interest of this research, the convergence test for MCMC will employ Geweke z-score test. In Geweke z-score test, convergence is determined by the difference of two means of the segments of the chain. Considering a chain having two segments, the initial segment A and segment B at the end of the chain. Taking chain A to be of length N_A and B has length N_B, then the two means are calculated as follows

$$:\begin{cases} \bar{\psi}_A = \frac{1}{N_A}\sum_{j=1}^{N_A}\psi_j \\ \bar{\psi}_B = \frac{1}{N_B}\sum_{j=N-N_B+1}^{N}\psi_j \end{cases}$$

Comparing the two means needs also an estimate of the sample variances ${\sigma_A}^2$ and ${\sigma_B}^2$. This task is complicated since the elements of the Markov chain are dependent. Therefore another estimator of variances is required. To ensure that $\bar{\psi}_A$ and $\bar{\psi}_B$ are independent, a 10% of iteration for the initial chain A ($N_A = {N}/{10}$) and 50% of iteration for last chain B ($N_B = {N}/{2}$) is taken. This is referred to as thinning of the chain and by doing this a distance between the two parts is created. If the ratios $\frac{N_A}{N}$ and is assumed to be fixed, then the two means are assumed to follow the same distribution and uncorrelated and Z-score follow Gaussian normal distribution, $Z_G \sim N(0,1)$.

Therefore:

$$Z_G = \frac{\bar{\psi}_B - \bar{\psi}_A}{\sqrt{\sigma_A{}^2} + \sqrt{\sigma_B{}^2}} \xrightarrow{d} (0,1)$$

This result is used to test the null hypothesis of equal location and is rejected if $|Z_G|$ is large indicating the chains have not converged.

2.10 Hierarchical Bayesian model

2.10.1 Introduction

In ordinary Bayesian regression, the hierarchical nature of the dataset and the source of random effect has not been considered. Assuming that TDHS 2015-16 dataset is a multi-layered dataset collected from complex sampling methods such as cluster sampling, this requires a model that takes into account these properties (Kamata, 2001; McCullagh & Nelder, 1989). Previously, data was assumed to be aggregated and the multi-layered nature was disregarded.

One important feature when units are drawn from clusters such as community, schools or households, is that it can no longer assume they are independent. It is becoming increasingly difficult to ignore the fact that units drawn from the same cluster are more similar to each other than they are units from the different cluster and the same applies to units from the same household(McCullagh & Nelder, 1989). Thus, the unobserved variables will induce correlation between units within the cluster which cannot be explained by the covariates within the model(Ng et al., 2006).

In recent years, researchers have shown Hierarchical Bayesian models to be a rigorous way to make scientific inferences about the population (Goldstein, 2010). In contrast to normal Bayesian regression where the data are assumed independent, however, in Hierarchical Bayesian models, data are not independent and therefore the need to compensate for the biases-

largely in standard errors which may arise when independence assumption is violated(Goldstein, 2010; Ng et al., 2006).

On the other hand, there is a difference in how the predictors affect the outcome of interest across the clusters and this is the main interest in employing the Hierarchical Bayesian model. This calls for a stepwise modelling approach for each parameter.

The multi-level nature of the dataset largely occurs when complex methods of data collection are used such as multistage cluster sampling results into multiple sources of variation. The practical instances in which one can encounter hierarchical data in the case of Tanzania may include an individual who is nested in a household, which is nested in a village, which is also nested in a ward and the ward nested in a district. The same practical implication is observed in the TDHS dataset whereby enumerated individuals are nested in households that are also nested in clusters. Taking into account the multi-layered nature of our dataset, it would be interesting to use the appropriate methods which take into account intra-class correlation (ICC). One needs to apply a model which introduces a random component that accounts for correlation among groups. A much more classical approach is Hierarchical logistic regression embedded in the Generalized Linear Mixed Model (GLMM) to capture multiple sources of variability (Kamata, 2001).

In this study, perinatal mortality exhibits variability within regions and areas of residences (McCulloch & Neuhaus, 2005). Tanzania is a large country with many administrative regions that exhibit a high degree of variability in terms of social-economic wealth index, maternal education and age at birth which may also impact variability in perinatal mortality. With all these differences, indeed the Generalized Linear mixed Model becomes instrumental for carrying out regional segmentation of the perinatal mortality (ECMS, 2016).

In Hierarchical models, there is a high probability of correlation of variance of residual errors between individual observations which are due to the nested nature of the data. The high correlation or dependence between observations makes traditional logistic regression inappropriate. The correlation or dependency in multi-stage cluster sampling occurs at several levels of the hierarchy (McCulloch & Neuhaus, 2005). For us to draw appropriate inferences, we require complex and tricky modelling techniques like Hierarchical modelling (Khan & Shaw, 2011; McCulloch & Neuhaus, 2005).

Hierarchical Bayesian modelling has gained a wide application currently due to its versatility in dealing with correlation among groups in the hierarchical data (Teacy et al., 2012). There is no number of studies that have used Hierarchical Bayesian modelling to solve different problems. Hierarchical Bayesian modelling was used by Teacy (Teacy et al., 2012) to study efficient and versatile approaches to trust and reputation and Andrade & Teixeira, 2015 to study the Railway track Geometry degradation which was very useful information for railway maintenance planners. It was also used to study geographic variability in age-dependent death rates(Arató et al., 2006). In a similar context, Zhu et al., 2006 applied Hierarchical Bayesian Spatial modelling to determine the availability of alcohol, drug hot spot and violent crime in Houston, Texas. The outcome suggested that activities around illicit drug outlets were more strongly associated with violent crime than alcohol outlets. The same approach was used to estimate the prevalence of HIV in Zimbabwe(Chinomona & Mwambi, 2015).

From the studies referred here, all have employed the Markov chain Monte Carlo algorithm for sampling from the posterior distribution. This has been done despite the complex nature of the datasets used(Andrade & Teixeira, 2015; Miklós Arató et al., 2006; Teacy et al., 2012; Zhu et al., 2006). Although extensive research has been carried out on Hierarchical Bayesian

modelling, very few studies have commercialized on Integrated Nested Laplace Approximation (INLA) due to its advantages over the MCMC. Therefore, this study is seeking to use the INLA method of inference to get the marginal posterior distribution as opposed to the MCMC approach.

2.10.2 Model Description

In this research, the aim is to estimate perinatal mortality from individuals who are also nested in households that are finally nested in clusters. This means perinatal mortality is likely to differ from different households and among clusters. On the other hand, perinatal mortality pattern is more likely to be correlated between members of the same household as well as for households within the same cluster. By aggregating and fitting a single linear model, assumes that deaths occurring in the various categories are independent. However, this is not the case since individuals in the same household are likely to have the same experience of perinatal mortality. Similarly, individuals from the same cluster are more likely to have the same experience on perinatal mortality than individuals from another cluster.

We may decide to model each household or cluster as a single model to satisfy the condition of the independent model. However, by so doing we will be throwing useful information by separating each model from each other. A better way to get rid of all these inconveniences is to employ a Hierarchical model structure to hierarchical data. The complexity which will be brought by this will create greater statistical power than isolated models.

2.10.3 Hierarchical Model for Binary Response Variable

If we consider our outcome of interest, perinatal mortality X_{ijk} to be binary response which takes values 1 if death occurs and 0 if no death, the choice of the model will be logistic regression as explained by McCullagh& Nelder(McCullagh & Nelder, 1989) and Kamata (Kamata, 2001). In

Generalized Linear Model (GLM) we can utilize one of several link functions including logit, probit and complementary log-log.

2.10.4 Generalized Linear Mixed Effects Model

Generalized Linear Mixed Effects Model (GLMM) is an extension of a class of Generalized Linear Model (GLM) in which random effects are added to the linear predictor's (Kamata, 2001; McCulloch & Neuhaus, 2005). The model assumes a predictor model with two portions i.e the fixed effect portion $X'\beta$ and the random effect portion $Z'\upsilon$ and it is the assumption for random effect model ($Z'\upsilon$) which induces correlation among observations. This is well illustrated by McCulloch(McCulloch & Neuhaus, 2005) that correlation coefficient $\text{corr}(Y_{ijk}, Y_{ijk}')$ depend on random effect. In this family of models we synthesize the three common widely models: The Generalized linear model (McCullagh & Nelder, 1989), mixed linear models having both fixed effect and random as well as models with structured dispersions. These types of models have the fixed effect and a random effect part to account for correlation between groups. For a multi-layered dataset as in our case, the random effect accounts for correlation in households and clusters.

In Bayesian parameter estimation, the non-informative priors and beta conjugate priors will be used. Since the outcome variable is Bernoulli, the beta conjugate prior is the prior of choice.

2.11 Prior Distribution

2.11.1 Informative Priors

Informative priors are the one that encodes useful domain knowledge about the parameters. The prior knowledge we intend to apply in this study is that perinatal mortality is 41.6 deaths per 1000 live births from the study by Mboya (Mboya et al., 2020).

2.11.2 Non-Informative Priors

Non-informative priors are used when nothing is known about the value of the parameter. These kinds of priors have minimal effect on the posterior distributions.

2.11.3 Conjugate Priors

Conjugate priors are priors with prior distribution and posterior distribution taking the same probability distribution.

2.12 Choice of Priors

In this study, we intend to use non-informative and beta conjugate priors to produce the marginal posterior distribution.

2.12 posterior distribution

In Bayesian analysis, posterior distribution about parameter θ represents the knowledge about parameter θ after having observed data x. This is often summarized as:

$f(\theta|x) = f(x|\theta)(f(\theta)$.

From the posterior distribution, we can estimate point location and credible regions. Point locations implicitly mean that a point statistic is used to summarize the posterior distribution (Pateras, K & Ntzoufras,I, 2013). The important point locations are mean, standard deviation, median, mode and quantiles(Chinomona & Mwambi, 2015). We can also estimate the quantiles and credible intervals of parameter and strictly speaking, the 100(1-q) credible interval is equivalent to 100(1-q) confidence interval for frequentist approach and they only differ in their interpretation (Wakefield, 2013). While confidence intervals are the measure of uncertainty around the effect estimate and they are based on sampling distributions, credible intervals can be interpreted as a 100(1-q) probability that the true

unobserved value would lie within the interval, given the evidence provided by the observed data (Hespanhol et al., 2019).

2.13 parameter estimation

The most common method for parameter estimation from the joint posterior distribution is the Markov chain Monte Carlo method (MCMC). However, this method is computationally very expensive and takes time in the computation process (Gómez-Rubio, 2020).

Despite the enormous impact of MCMC on statistical inference, they can be computationally burdensome, especially for complex models such as hierarchical model (Taylor & Diggle, 2014). On top of this, the inferential validity rests on the convergence of the Markov chain distribution which becomes difficult to verify empirically and takes more time for hierarchical datasets with random effects (De Smedt et al., 2015; Taylor & Diggle, 2014). A novel method for inference known as Integrated Nested Laplace Approximation (INLA) will be used for the computation of the individual posterior marginals. This is because this method allows fast and accurate model fitting. A combination of analytical approximation and numerical integration in INLA circumvents the convergence issues hence making the process quicker (Taylor & Diggle, 2014). Despite these superiority properties, it has some disadvantages that the analytical approximation can introduce errors and when the event of interest is rare, it produces worse estimates and credible intervals(De Smedt et al., 2015; Taylor & Diggle, 2014).

2.13.1 Integrated Nested Laplace Approximation

Integrated Nested Laplace Approximation (INLA) was proposed by Havard Rue, Martino and Chopin(Rue et al., 2009). In contrast to the Markov chain Monte Carlo method which aims at estimating the joint posterior

probability, INLA aims at estimating the individual posterior marginals of the model parameter (Gómez-Rubio, 2020). If we consider the computation of the posterior mean $(\bar{\theta}) = \int \theta_{ip}(\theta_i|y)d\theta_i$ and the median $=\int \bar{\theta}_M\, p(\theta_i|y)d\theta_i$, we see that estimation of the location of posterior marginals aims at estimating the integrals that appear in Bayesian inference. In the equation for computation of the mean or median above, the term $\boldsymbol{p(\theta_i|y)}$ is the marginal posterior distribution of the univariate parameter $\boldsymbol{\theta_i}$.

2.13.2 Model choice and assessment

Similar to the Markov chain Monte Carlo (MCMC) method, INLA also computes several Bayesian criteria for the model assessment (Schrödle et al., 2011). The popular Deviance Information Criteria (DIC) was used for model choice and assessment.

The models included the full model with all predictor variables and models including subsets of full models. The DIC takes into account the model-fit and penalty term which originated from the complexity of the model via the estimated number of effective parameters (Gómez-Rubio, 2020; Schrödle et al., 2011). It is computed as the sum of the posterior mean of the deviance and the number of effective parameters.

$$DIC = D(\bar{x}, \bar{\theta}) + 2p_D$$

Generally, low mean deviance indicates good model fit while large mean deviance indicates poor fit. However additional effective parameters reduce the DIC making poor model fit. Therefore, in essence, the addition of several effective parameters is to penalize model complexity (Schrödle et al., 2011).

CHAPTER THREE: METHODOLOGY

3.1 Research area

This research was conducted in Tanzania using Tanzania demographic Health survey 2015-2016. Tanzania is a country located in East Africa and it is bordering Kenya to the North-East, Uganda to the North, the Indian Ocean to the East, the Democratic Republic of Congo to the West, Burundi and Rwanda to the North-West, Zambia to the South-West and Mozambique to the South. The country is located within **Latitude:** -6° 22' 22.17" S and **Longitude** 34° 53' 32.94" E. It is estimated to have a population of about 57.3 million according to the National Bureau of Statistics (2017).

The National Bureau of Statistics in collaboration with the Ministry of Health, Gender, Community Development, Elderly and Children jointly conducted TDHS 2015-2016 with significant technical and logistic assistance from the Ministry. The international stakeholders such as Demographic Health Program from the USA provided technical assistance throughout the project. The survey was a nationally representative sample with a probability sample of 13,376 households.

3.2 Research design

This is a retrospective study that was conducted in 2015-16 collecting data for five years back to 2010.

3.3 Study population

This is secondary data targeting the women population of age between 15-49 years from sampled households. The sample was subdivided into 59 strata and 608 clusters. Among 13,376 households, 12,563 households were successfully interviewed yielding a response rate of 98%. Of the clusters surveyed, 180 (29.6%) were from urban and 428 (70.4%) were from rural.

With an average of 22 households per cluster, 13,376 households were involved in the survey whereby 3,960 were from urban areas and 9,416 households were from rural areas.

3.4 Data Structure

The Tanzania Demographic Health Survey (TDHS) data is divided into 59 strata, 608 clusters which were regarded as enumeration areas (EAs) which in turn provided 22 households each. The country was divided into 9 regions for ease of operation namely: Western Zone, Northern Zone, Central Zone, Southern Highlands Zone, Southern Zone, Lake zone, South West Highlands Zone, Eastern zone and Zanzibar.

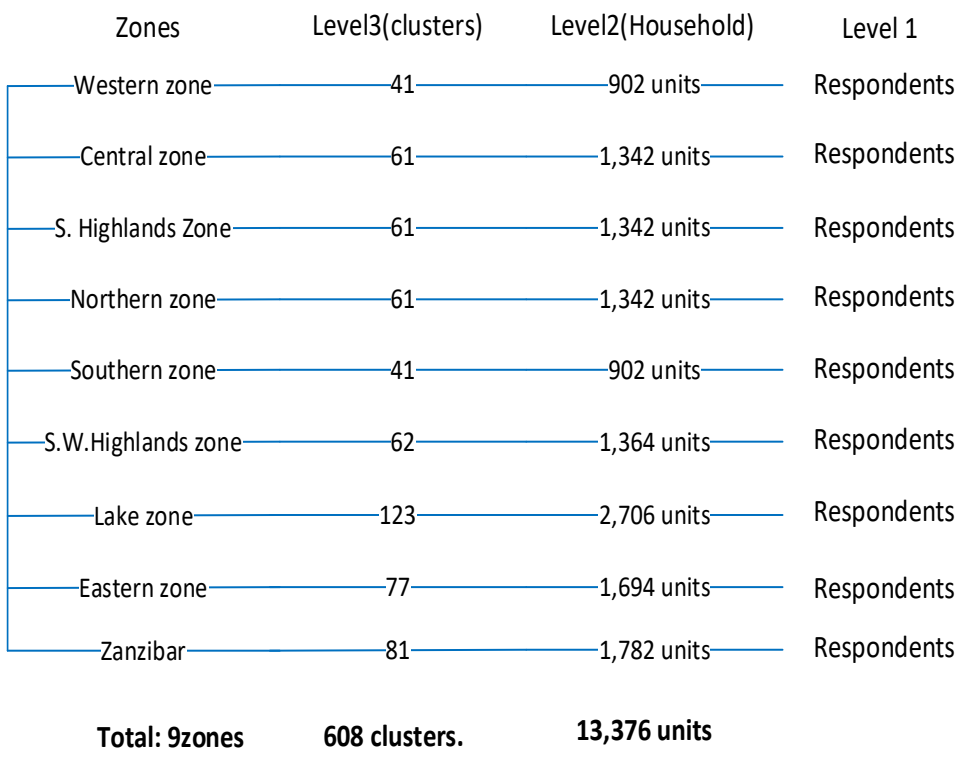


Figure 2: Hierarchical structure for 2015-16 TDHS dataset.

3.5 Study variables and their method of measurement

This study is aiming at finding a statistical association between perinatal mortality and a set of independent variables.

3.5.1 Dependent variable

Perinatal mortality includes stillbirth and neonatal deaths during the first week of life. In this study the dependent variable is a binary outcome (perinatal mortality) which can take two forms i.e., death occurring or not.

Therefore: $Perinatal\ mortality = \begin{cases} 1\ if\ death\ occur \\ 0\ if\ no\ death \end{cases}$

3.5.2 Independent variables

The factors we want to model for perinatal mortality in Tanzania are maternal demographic characteristics such as mother's age at first birth, marital status, maternal weight, maternal BMI, and pregnancy interval. Other variables are social-economic characteristics such as place of residence, region of residence, maternal education and social-economic wealth index. Others include birth order, child's sex, marital status and region of residence. The dummy variables for each independent variable were created.

3.6 Data management

Data was requested from the Demographic Health survey DHS program 530 Gaither Road, Suite 500, Rockville, MD20850 USA. Data files in Stata format were kept in a secure file on a personal computer (Macbook air®). The variables were selected mainly from the literature background. The selected variables were categorized based on the criteria obtained from the DHS report. The independent variables were coded with numbers and dummy variables created for each variable say for example for maternal

age: 1 representing <20 years, 2 representing age 20-29 years, 3 representing 30-39 years and 4 representing 40-49 years

3.7 Conceptual framework

Generally, there are multidimensional interactions of community factors, socio-economic factors and proximate factors which can influence perinatal mortality in the community (Biracyaza E* and Habimana S, 2019). In conceptualization, this study will use the framework developed by Mosley and Chen(Mosley & Chen, 2003) that presents events in life that can result in perinatal mortality. However, the framework is modified to include other determinants which cause perinatal mortality during pregnancy.

The framework incorporates both social and biological determinants of perinatal mortality and also integrates research methods employed by social and medical scientists (Mosley & Chen, 2003). The interaction between different factors for perinatal mortality is depicted in *figure 3.*

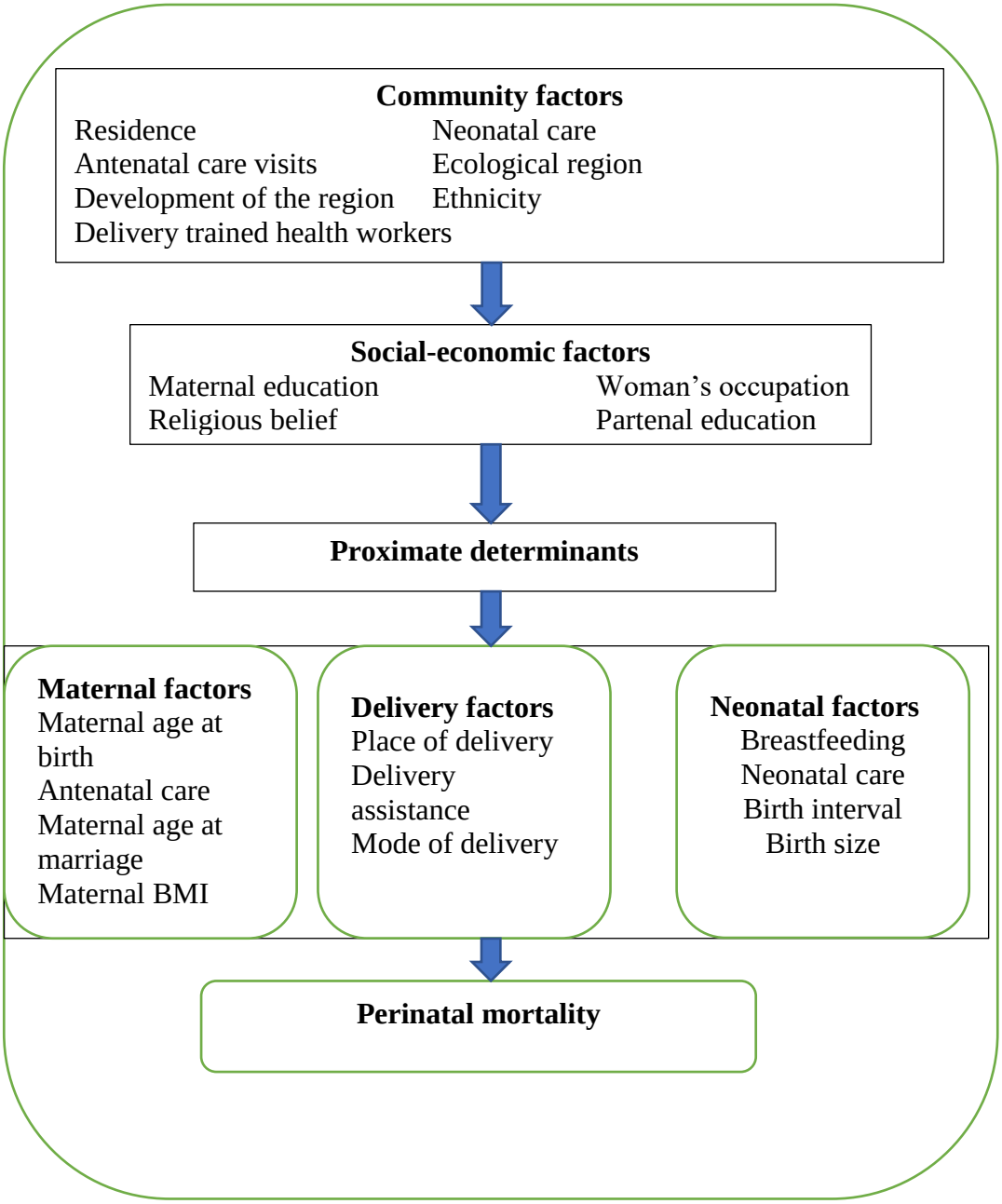


Figure 3: Model interaction for the community, social-economic and proximate Factors

3.8 Ethical considerations

Ethical approval was obtained from Institutional Review Board, Kenyatta

The hospital and the permission to access data was given by the DHS program.

3.9 Statistical Analysis

3.*9.1* Descriptive statistics

We computed the distribution of perinatal mortality rates among the selected explanatory variables. Perinatal mortality was design-based and computed based on the number of live births per respective level or class. The weighted number of stillbirths and neonatal deaths were computed in each variable category as well as the prevalence in percentages. The sum of stillbirths and neonatal deaths is divided by the number of live births provided perinatal mortality.

3.9.2. Bivariate chi-square test

The statistical association between individual variables and perinatal mortality as an outcome variable was checked using a bivariate chi-square test.

3.*9.3* Multicollinearity test

Multicollinearity is a phenomenon whereby one or two predictor variables are correlated. In regression analysis, we seek to find the correlation between the outcome variable and a set of predictor variables. In some cases, there may be a correlation among predictors of something undesirable. If this happens, the standard error of the coefficients will increase making some variables statistically insignificant while they are significant hence affecting the reliability of the regression model. When collinearity exists, the standard error of the predictor's coefficients will increase and ultimately the variance of predictor's coefficients will be inflated. One of the tools to test this is the variance inflation factor (VIF).

The interpretation is based on the following criteria: VIF=1 no correlation, 1<VIF<5 moderately correlated, VIF>5 highly correlated.

3.9.4 Variable selection

The stepAIC function in package MASS(Venables & Ripley., 2002) was used for both backward and forward stepwise variable selection. All twelve variables were fitted in the model. This is because in combination with other variables they can explain the dependent variable rather than when they are alone. The bivariate test outcome was also considered for the variables to be included in the model.

3.9.5 Bayesian Analysis

Statistical analysis was performed using R version 4.0.3(R Core Team (2014), n.d.). For ordinary Bayesian analysis, the data was aggregated and assumed independent and Bayesian logistic regression was used. The credible interval chosen for this study was 95%. The DIC was used for model choice and assessment. The choice of the best fit was done using the package BayesFactor (Morey & Rouder, 2018) using function regressionBF.

In the case of women who had a stillbirth and those who had their neonates dying within the first 7 days of life, for woman number I pos[i] is the binary indicator that a woman had perinatal death or not (coded 1 if a woman had perinatal death and 0 if not). Therefore, the data were analyzed using the Bayesian logistic regression model:

$$\text{top} Y_i \sim \text{Bernouli}(q_i) \text{ and logit } (q_i) = \sum_{j=1}^{p} X_{ijk}\, \beta_j$$

Whereby q_i is the binary outcome (perinatal mortality), X_{ijk} is a vector of covariate matrix from the i[th] woman from j[th] household and k^{th} cluster and β_j a vector of regression coefficients.

Given our prior distribution we can have our generative model as follows:

$$logitq_i = \beta_{0j} + \beta_1 \times Residence + \beta_2 \times Mateducat + \beta_3 \times BirthInt + \beta_4 \times Mateage + \beta_5 \times SEWI \ldots\ldots\ldots \beta_n X_n$$

Having established the priors, dataset and generative model, the Markov Chain Monte Carlo (MCMC) algorithm was used to find the posterior distribution of the model parameters.

Bayes factor was used for the selection of the variables or discrimination of other competing models. Given two models, M1 and M2, Bayes factor for comparing the two models M1 and M2 are numbers $\beta_{\mathbf{1,2}}$ such that:

$$\mathrm{p(M1\,|\,y)}\Big/_{\mathrm{p(M2\,|\,y)}} = \beta_{1,2}\, p(M1)\Big/_{p(M2),}$$

Where $\boldsymbol{p(M1)}$ and $\boldsymbol{p(M2)}$ indicate the prior odds for each model (Souza & Migon, 2004). The Bayes factor is a multiplier that changes the prior odds for the model into the posterior odds. The convergence test was performed using Geweke z-score test.

3.9.6 Hierarchical Bayesian Analysis

Following the fundamental properties of Hierarchical modelling and the structure of the dataset, a three-level equation can be generated according to the generalized linear model framework as follows:

$$g\left[E\left(X_{ijk}\right)\right] = \alpha_{ijk} + \beta_{jk} + \gamma_k$$

Where $\boldsymbol{g}(*)$ is a link function and $\alpha_{ijk} + \beta_{jk} + \gamma_k$ implies respondent, household and cluster respectively.

Since we are dealing with a dataset with three levels i.e., cluster, household and individual respondent we assume that the measurement on the response variable can be denoted as X_{ijk} expressing i^{th} individual in the j^{th} household of the k^{th} cluster. At i^{th} person level, the coefficients β_s are not constant across all persons. Therefore, we can break up the three levels and set up a linear model at each level of the hierarchy based on the generalized linear model (GLM) framework by introducing a random variable that explains the difference among clusters, households and respondents.

Each generalized linear model of each hierarchy is a mixed effect model. If we combine the three linear models for each hierarchy, we can produce a full generalized linear model which represents individual respondent, house level and cluster level as follows:

$$X_{ijk} = \gamma_0 + \alpha_0 + \beta_0 + \sum_{l=0}^{q} \gamma_l \, w_{l,k} + \sum_{l=0}^{p} \beta_{l,k} \, z_{l,jk} + \sum_{l=0}^{r} \alpha_{l,jk} \, x_{l,ijk} + \epsilon_k + \varepsilon_{jk} + \kappa_{ijk}$$

Where ϵ_k is a random variable with $E(\epsilon_k) = 0$ and var(ϵ_k)=σ_ϵ^2, ε_{jk} is a random variable with $\mathrm{E}(\varepsilon_{jk}) = 0$ and $\mathrm{var}(\varepsilon_{jk}) = \sigma_\varepsilon^2$, κ_{ijk} is a random variable with $E(\kappa_{ijk}) = 0$ and var(κ_{ijk})=σ^2. Since the individual model for each hierarchy is a mixed effect model, the combined full model is also a mixed effect model. The above model is a benchmark for hierarchical Bayesian analysis.

In Hierarchical Bayesian analysis, data were not aggregated and correlation was assumed. Integrated Nested Laplace Approximation (INLA) method was used for parameter estimation to determine the association between covariates and perinatal mortality. Considering y_i to be a binary response variable in this case perinatal mortality, it takes a value of 1 if death occurs and 0 if death does not occur.

$y \sim Bernoulli(x_i, p_i)$

Where $x_i = \{x_1, x_2, \ldots . x_p\}$ is a vector of covariates such as maternal education, age at birth etc. and p_i is a proportional of neonates died. From the latent model with structured additive predictors which is defined as:

$\boldsymbol{\eta_i = \alpha + \sum_{j=1}^{i} \beta_j Z_{ji} + \sum_{k=1}^{i} f^k (v_{ki}) + \varepsilon_{ij}}$ whereby:

$\boldsymbol{\alpha}$ is the intercept, $\beta\, is$ the coefficients of covariates and f^k is a function that describes the random effect on some vectors of covariates and ε_{ji} is an error term.

Bayesian structured additive logistic regression was preferred in this case because of its ease of interpretation. Different models ranging from the full model and models containing subsets of the covariates were compared by using deviance information criteria (DIC).

The model fit was assessed using DIC and the model with the lowest DIC was considered a good fit. From marginal posterior distribution, we can compute the posterior mode denoted as $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}}$ defined as $\mathbf{arg}\, \boldsymbol{max_p}\, \boldsymbol{p(\theta|y)}$ which provides the value of $\boldsymbol{\theta}$ when the value of $\boldsymbol{p(\theta|y)}$ is maximal while posterior mean $\overline{\boldsymbol{\theta}} = \boldsymbol{E(\theta_i|y)} = \int \boldsymbol{\theta_{ip}(\theta_i|y)d\theta_i}$ and the optimal choice leads to minimizing the expected loss. The posterior median $\boldsymbol{\theta_i(0.5)}$ is often considered the adequate summary of the location of posterior marginal distribution (Wakefield, 2013) and is always the solution to equation $0.5 = \int \overline{\boldsymbol{\theta}}_{\boldsymbol{M}}\, \boldsymbol{p(\theta_i|y)d\theta_i}$.

We could also estimate the quantiles and credible intervals of parameter $\boldsymbol{\theta_i}$ as follows:

The 100*q% quantile, $\boldsymbol{\theta_i(q)}$ with 0<q<1 is found by solving

$Q = \Pr[q = \Pr(\theta_i \leq \theta_i(q)] = \int_{-\infty}^{\theta_i(q)} p(\theta_i\,|y) d\theta_i$

The $\mathbf{100(1-q)}$ equal-tailed credible interval whereby $\mathbf{0 < q < 1}$ is given by:

$$[\theta_i(1-q)/2, \theta_i(1+q)/2)]$$

If we consider a vector of $\boldsymbol{n}$ observations $\boldsymbol{y} = \{\boldsymbol{y_1}, \boldsymbol{y_2} \ldots . \boldsymbol{y_n}\}$ we can establish a convenient generalized linear model with link $\boldsymbol{\eta}$ as follows:

$\boldsymbol{\eta_i} = \boldsymbol{\alpha} + \sum_{\boldsymbol{j=1}}^{\boldsymbol{i}} \boldsymbol{\beta_j Z_{ji}} + \sum_{\boldsymbol{k=1}}^{\boldsymbol{i}} \boldsymbol{f^k}(\boldsymbol{v_{ki}}) + \boldsymbol{\varepsilon_{ij}}\ \boldsymbol{i} = \boldsymbol{1, 2} \ldots, \boldsymbol{n}$ whereby

$\boldsymbol{\alpha}$ is the intercept, $\boldsymbol{\beta}\ is$ the coefficients of covariates such that $\{\boldsymbol{z_j}\}_{\boldsymbol{j=1}}^{\boldsymbol{i}}$ and $\boldsymbol{f^k}$ is a function that describes the random effect on some vectors of covariates $\{\boldsymbol{v_{ki}}\}_{\boldsymbol{k=1}}^{\boldsymbol{i}}$ and $\boldsymbol{\varepsilon_{ji}}$ is an error term. Since this model can estimate parameters such that $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{f}(.),\}$ it is called latent model(Musio et al., 2013; Schrödle et al., 2011). In integrated Nested Laplace Approximation (INLA) the vector of latent effect $\boldsymbol{x}$ is given as

$$x = (\eta_1, \ldots . \eta_n, \alpha, \beta_1 \ldots .)$$

The latent structure is considered as Gaussian Markov Random Field (GMRF) since it has Gaussian properties and it is Gaussian random variable. Therefore, the latent field can be estimated by a Gaussian Markov Random Field as:

$\theta = \{\alpha, \beta, \eta, f(.),\} \sim GMRF(\psi)$

were

θ is a set of parameters (latent field)

α is the intercept parameter

β the linear predictor effect parameters (coefficients)

f^k the set of k functions associated with non-linear covariates

ψ are hyperparameters representing variance-covariance matrix, Σ.

As it was discussed earlier, the component of random effect is assumed to be drawn from a multivariate normal distribution with mean 0 and variance-covariance matrix, Σ i.e $\mathrm{N}(0, \Sigma)$. For models which assume independence, the off-diagonals are assumed zeros

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}$$

However, in many situations, there are correlations or random effect dependency where the off-diagonals are not all zeros and instead, there is a high degree of dependency among observations. The variance-covariance matrix is given as:

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho & \rho^3 \\ \rho & \sigma^2 & \rho^2 \\ \rho^3 & \rho^2 & \sigma^2 \end{pmatrix}$$

It should be noted that there is an exponential increase in covariance with the increase in a number of times (n). Since the calculation of covariances involves multiplication of a variance-covariance matrix by the inverse of covariance matrix which then involves substantial expansion which rapidly bursts and exhaust the memory. To avoid all these, we implement the use of the Gaussian Markov Random Field (GMRF)

The essence of combining all the parameters into θ is to avoid computing each parameter individually. Given the structure of GMRF, the precision of θ will be very sparse and INLA takes advantage of this sparse structure and conditional independence of GMRF to speed the computation (Gómez-Rubio, 2020). Now if we consider a vector of latent effect, x and vector of latent field θ we can have the following equation:

$$\pi(\theta, x|y) = \frac{\pi(y|\theta, x)\pi(\theta, x)}{\pi(y)} \propto \pi(y|\theta, x)\pi(\theta, x)$$

In this equation $\pi(y)$ is the marginal likelihood or normalizing constant which is difficult to compute and most of the time it is ignored from the equation. Also $\pi(y|\theta, x)$ is the likelihood.

From the latent effect x and θ we can compute the likelihood as follows:

$$(y|\theta, x) = \prod_{i \in I} \pi(y|x_i, \theta)$$

We know that the intention of INLA is not to compute the posterior distribution but to compute marginal posterior distribution for the latent effect x and hyperparameters θ. For hyperparameters xl we can compute

$$\pi(xl|y) = \int \underbrace{\pi(xl,\theta|y)}_{part\ 2}\underbrace{\pi(\theta|y)}_{part\ 1}d\theta$$

Part 1 and 2 in the above equation can be processed to give a very useful meaning such that $\pi(x,\theta,y) = \pi(x|\theta,y)\pi(\theta|y)\pi(y)$.

Part 1 of the integral can be estimated by:

$\tilde{\pi}(\theta|y) = \frac{\pi(x,\theta,y)}{\tilde{\pi}_G(x|\theta,y)}\Big| x = x^*(0)$ which is the Laplace approximation of marginal posterior distribution and the factor $\tilde{\pi}_G(x|\theta,y)$ represent the Gaussian distribution to $\pi(x|\theta,y)$ and $x^*(0)$ is the mode of the full conditional of x for a given θ

We can also compute marginal posterior for θ in a similar way:

$\pi(\theta|y) = \int \pi(\theta|y)d\theta_{-k}$ where θ_{-k} is a vector of hyperparameters θ without θ_k.

From the previous knowledge that perinatal mortality is 41.6 deaths per 1000 live births, it can be established from $(x+1, n-x+1)$ the hyperpriors will be $Beta(\alpha = 42.6, \beta = 959.4)$ assuming perinatal mortality in Tanzania is 41.6(Mboya et al., 2020)

3.9.7 Likelihood of The Model

From the knowledge that GLMM incorporates both fixed and random effects, then a model with an n-dimensional vector with linear predictors can be expressed as follows:

$$E[X_{ijk}|\mathrm{X}', \upsilon] = g^{-}(X'\beta + Z'\upsilon)$$

Where $\boldsymbol{Z'}$ is an nxq model matrix of covariates with random effects and $\boldsymbol{g}$ is a link function (e.g., logit, probit, log). The term $\boldsymbol{X'\beta}$ is a fixed effect term. In the Bayesian world, all the parameters are assumed to be random variables and there is no need to partition the equation into two. In this

expression, $\boldsymbol{v}$ is a component of random effect which is associated with the level of grouping and is generally assumed to follow normal multivariate distribution $\boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ with density function $\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{\Sigma})$ (Chinomona & Mwambi, 2015; Ng et al., 2006).

This means:

$\boldsymbol{v} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$

Now, we can maximize the marginal likelihood of the observed data using the estimates β and Σ to get the marginal density of Y_{ijk} given β and Σ. The likelihood for k observations is given by:

$L(X_{ijk}|\beta, \Sigma) = \int \prod_{i=1}^{k} \{p(X_{ijk}|\upsilon, \beta, X', Z')\} f(\upsilon|\Sigma) du$

Whereas $p(X_{ijk}|\upsilon, \beta)$ is the probability density function of X_{ijk} given υ and β and $f(\upsilon|\Sigma)$ is the probability density function of υ given Σ.

As it was discussed by Edmond SW Ng (Ng et al., 2006), the likelihood could be easily tractable if the response model was normal and the link $\boldsymbol{g}$ identity. Therefore, since the model response is non-normal, the maximum likelihood estimates are intractable and can be found through computing methods such as Monte Carlo integration or Integrated Nested Laplace Approximation (INLA)

3.9.8 Study Dissemination Plan

The study results will be disseminated using publication to a reputable journal. Also, the study results will be shared with Demographic Health Survey (DHS) program.

3.9.9 Study Limitations

There is a relatively large number of stillbirth and neonatal deaths that goes unreported. This is attributed to many factors such as recall bias, communication barriers, delivery at homes etc. In this study, this is dealt

with by applying the Bayesian approach which can enable us to apply prior beliefs on regression coefficients for each covariate.

CHAPTER FOUR: RESULTS

4.1 Descriptive Results

This study involved a total of 13,266 women aged between 15-49 years and 10,052 children born in the previous 5 years period. The weighted and unweighted number of stillbirths were calculated from children files while weighted and unweighted numbers of neonatal deaths were computed from women files. From the two files, the number of stillbirths was 187 while neonatal deaths were 215. This depicted perinatal mortality of 402 equivalent to the one reported in the TDHS 2015-16 report (Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) & ICF, 2016).

Table 2 presents the results of the computed perinatal mortality per 1000 live births and bivariate statistical association using the chi-square test. It was noted that perinatal mortality was higher in women aged <20 years and older women aged between 40-49 years compared to the mid-ages such as 20-29 and 30-39 and the difference was statistically significant($p<0.0001$). There were rural-urban differences where mortality was higher in urban residence (48 deaths per 1000 live births) and 37 deaths per 1000 live births in rural residences. However, the difference was not significant ($p<0.2846$).

Also, of interest, perinatal mortality was higher in the first pregnancy and decreased with proceeding birth orders ($p<0.0001$) while it was also higher in <15 months birth interval, decreasing progressively with the increase of birth interval ($p<0.0001$). Perinatal mortality was relatively low in women with higher education and relatively higher in women with lower education i.e., women with no education, incomplete primary education and complete primary education ($p<0.0001$). Perinatal mortality in women who never

married was relatively higher compared to other marital statuses of women but the difference was not significant (p=0.4146). It was also observed that perinatal mortality was relatively higher in women with <20 years (59.6 per 1000) and decreased with an increase in age. However, there was a slight increase in mortality in higher ages i.e., 41.3 per 1000 in age 40-49. Again, those women who got their first babies aged 30-39 years experienced higher perinatal mortality rates than other women. Both maternal ages and maternal age at first birth were statistically significant (p<0.0001).

Table 1: Maternal and child variable distribution and bivariate association.

Variables	Weighted (n%)	Stillbirth (n%)	Neonatal Deaths (n%)	Perinatal Deaths (/1000)	p-value
Residence					
Rural	9121 (68.8)	44(23.5)	87 (40.5)	48	0.28
Urban	4145 (31.2)	143(76.5)	128.(59.5)	37	
Wealth Index					
poorest	2144 (16.2)	38 (20.3)	37 (17.2)	31	
poorer	2166 (16.3)	40 (21.4)	47 (21.9)	41	
middle	2438 (18.4)	52 (27.8)	32 (14.9)	43	0.03
rich	3108 (23.4)	33 (17.6)	52 (24.2)	45	
richest	3410 (25.7)	25 (13.4)	48 (22.3)	44	
Education					
No education	1998 (15.1)	42 (22.5)	21 (9.8)	30	<0.0001
Incomplete Primary	7640 (57.6)	125(66.8)	161(74.9)	44	
Complete Primary	3487 (26.3)	19 (10.2)	32 (14.9)	38	
Secondaryor Higher	141 (1.1)	1 (0.5)	1 (0.5)	23	
Occupation					
Not working	3773 (28.4)	29 (15.5)	50 (23.3)	36.7	<0.0001
Working	9489 (71.5)	158(84.5)	165 (76.7)	40.8	
Missing	4 (0.03)				
Marital status					
Never married	3478 (26.2)	38 (20.3)	37 (17.2)	31	
Married	6137 (46.3)	40 (21.4)	47 (21.9)	41	0.41
Living together	2052 (15.5)	52 (27.8)	32 (14.9)	43	
Divorced	1254 (9.5)	33 (17.6)	52 (24.2)	45	
Widowed	345 (2.6)	25 (13.4)	48 (22.3)	44	
Weight(kg)					
<50	3,799(28.6)	53(28.3)	38(17.7)	34.5	0.28
50-70	7,475(56.3)	105(56.1)	126(58.6)	38	
70-90	1,584(11.9)	23(12.3)	47(21.9)	63.6	
90-110	287 (2.2)	3(1.6)	3(1.4)	40	
>110	121 (0.9)	3(1.6)	2(0.9)	61	
Body Mass Index					
<18.5	3,799(28.6)	7 (3.7)	5 (2.3)	18	
18.5-24.9	7,475(56.3)	123(65.8)	131 (60.9)	37.6	0.20

25-29.9	1,584(11.9)	44(23.5)	44 (20.5)	50.5	
>30	287 (2.2)	13 (7.0)	35 (16.3)	61	
Missing	129 (1)				
Age at first birth					
<20	6973(52.6)	9(48.7)	136 (63.3)	36.2	
20-29	2,603(19.6)	67 (35.8)	75 (34.9)	38.8	
30-39	144 (1.1)	3 (1.6)	4 (1.9)	56	<0.0001
40-49	1 (0.0)	0 (0.0)	0 (0.0)	0	
Missing	3,545(26.7)	26(13.9)	0 (0.0)		
Age (years)					
<20	2,932(22.1)	15(8.0)	26 (12.1)	59.6	
20-29	4,577(34.5)	96(51.3)	97 (45.1)	38.8	<0.0001
30-39	3,375(25.4)	55(29.4)	70 (32.6)	37	
40-49	2,382(18.0)	21(11.2)	21 (9.8)	41.3	
Total	**13,266**				
Previous birth interval					
<15	285(2.8)	13 (7.0)	34 (15.8)	33	
15-26	2063(20.2)	33(17.6)	32 (14.9)	25.6	
26-38	2476(24.2)	23(12.3)	27 (12.6)	34	<0.0001
39+	2988(29.2)	43(23.0)	43 (20.0)	40	
Missing	2,421(23.7)				
Birth order					
1	2,399(23.4)	49 (26.2)	79 (36.7)	51	<0.0001
2-3	3,388(33.1)	47 (25.1)	56 (26.0)	30	
4-5	2,216(21.7)	34 (18.2)	35 (16.3)	32.5	
6+	2,230(21.8)	30 (16.0)	45 (20.9)	37	
Missing	0 (0)	27 (14.4)	-	-	
Gender					
Female	5,098(49.8)	83 (44.4)	132(61.4)	42	0.77
Male	4,954(48.4)	77 (41.2)	83 (38.6)	32.3	
Missing	181 (1.8)	27 (14.4)	-		
Total	**10,233**				

The prevalence of perinatal mortality among the variables greatly varied across all the variables. *Figure 4* shows the highest mortality occurred in obese women as well as heavyweight heavyweight women (>110kgs).

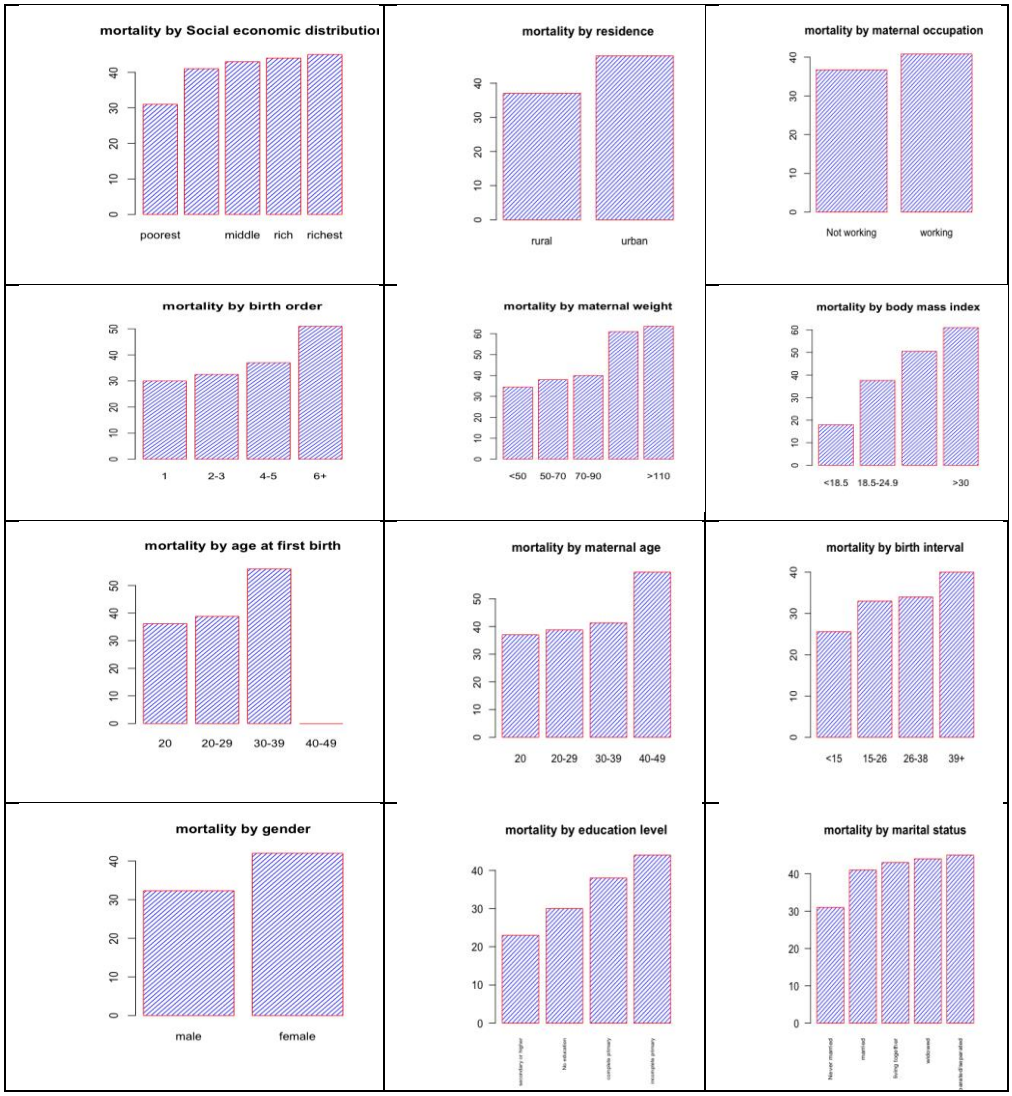


Figure 4: Prevalence of perinatal mortality among variables

4.2 Multicollinearity test

The multicollinearity test indicated a moderate correlation among the predictors. This is thought not to seriously affect the standard errors of the

predictor's coefficients. The variance inflation factor (VIF) for the predictor variables ranged between 1.15 for gender and 4.63 for variable birth order. Since the range was within the limits, then there was no multicollinearity among the variables.

4.3 NORMAL BAYESIAN ANALYSIS

4.3.1 Fitting the best model

The best model was determined using the Bayes factor and Bayesian information criteria (BIC) in which the model with the highest Bayes factor was considered the best model. Consequently, the model with the lowest BIC was considered the best model.

From twelve predictor variables, only two variables were statistically significant to explain perinatal mortality. The Bayes Factor method produced the best model containing birth order and previous birth interval. Consequently, the omission of these variables from the model reduced the Bayes factor to almost zero showing that the two variables were all significant. Secondly, the model with only birth order and previous birth interval had the minimum BIC (BIC=-3136.18) while the full model had the highest BIC (BIC=-3057.44). Therefore, the best fit for the model explaining perinatal mortality included only two variables namely previous pregnancy interval and birth order of a child. However, from literature and stepAIC function for variable selection other four variables and the interaction were included in the model.

4.3.2 Bayesian analysis using non-informative priors

4.3.2.1 Convergence test and diagnosis

Bayesian logistic regression analysis was computed using package arm (Gelman & Yu-Sung Su, 2020) in R. Convergence of the chains was

diagnosed using the Geweke convergence test. Convergence at MCMC algorithm was aimed at checking how close the process is to the true posterior distribution (Plummer et al., 2005). Also, the convergence diagnostic test was to check for stationarity of the process and verification of posterior summary measures. The Geweke test examines if the z-score falls within two standard deviation bands for the tested parameters (Gelman et al., 2013; Plummer et al., 2005). In this research, the intercept and all six variables were within the two-bound standard deviation indicating that convergence was attained at 50,000 iterations as is shown in *figure 5*.

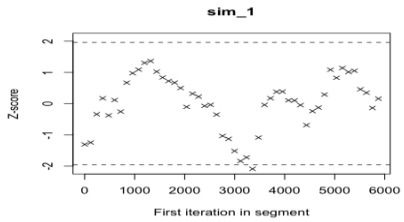
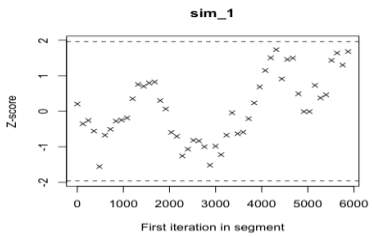
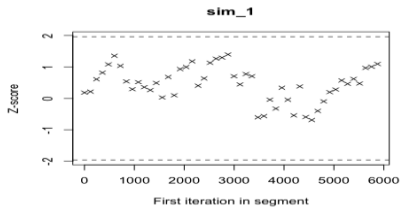
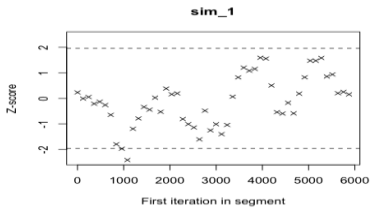
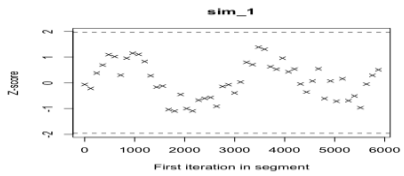
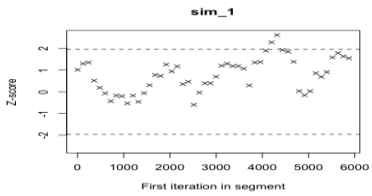
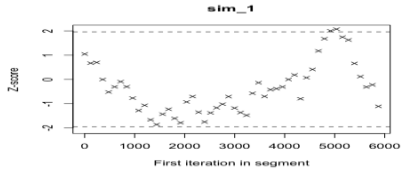
<table>
<tr>
<td>a) Intercept
 
</td>
<td>b) Birth order
 
</td>
</tr>
<tr>
<td>c) Previous birth interval
 
</td>
<td>d) Maternal age
 
</td>
</tr>
<tr>
<td>e) Occupation
 
</td>
<td>f) Maternal education
 
</td>
</tr>
<tr>
<td>g) Age at first birth
 
</td>
<td></td>
</tr>
</table>

Figure 5: Geweke plot for the convergence test.

Table 2 presents the results for Bayesian logistic regression whereby the three predictors including birth order, previous birth interval and maternal age were statistically significant since the 95% credible intervals for adjusted odds ratio (AOR) did not include 1 i.e., AOR=8.9(95% CI,5.81,13.74), AOR=7.76(95% CI,4.81,12.43), AOR=9.58(95% CI,5.53,16.61) for birth order, AOR=0.24(95% CI,0.17,0.36), AOR=0.18(95% CI,0.12,0.27), AOR=0.18(95% CI,0.12,0.20) for previous birth interval and AOR=3.42(95% CI,2.10,5.58), AOR=3.11(95% CI,1.75,5.47) and AOR=2.17(95% CI,1.16,4.10) for maternal age. The results also show that previous birth interval was associated with a lower risk of perinatal mortality while birth order and maternal age were associated with a high risk of perinatal mortality. Although maternal education was not statistically associated with perinatal mortality, it was associated with a lower risk of perinatal mortality.

Considering only the significant variables, the odds of perinatal mortality was 8.9 times more likely in 2-3 birth order compared to first birth, 7.76 times more likely in 4-5 birth order compared to first birth and 9.58 times more likely in 6+ birth order compared to first birth. The risk of perinatal mortality increased from 2+ birth orders. The odds of perinatal mortality were 76% less likely for birth interval 15-26 months compared to <15 months, 82% less likely for interval 26-38 months compared to <15 months and 82% less likely for 39+ months compared to <15 months. The risk of perinatal mortality decreased from 15+ months intervals. On the other hand, the odds of perinatal mortality were 3.42 times more likely for age 20-29 years compared to <20 years, 3.11 times more likely for 30-39 years and 2.17 times more likely for 40-49 years. The trend shows that the risk of perinatal mortality decreased from 20+ years of age. The interaction between occupation and maternal education was not statistically significant

which implies that the association between maternal age, maternal age at first birth, occupation, maternal education, birth order and previous pregnancy interval with perinatal mortality is not modified with interaction (effect modifier).

Table 2: Bayesian logistic regression output for the best fit

Variable	Estimate	Coefficient Standard error	AOR	AOR (95% CI)
Intercept	-4.83	0.36	0.008	(0.00, 0.02)
Birth order				
1(Ref)				
2-3	2.19	0.22	8.90	(5.81,13.74)
4-5	2.05	0.24	7.76	(4.81,12.43)
6+	2.26	0.28	9.58	(5.53,16.61)
Previous interval				
<15 months (Ref)				
15-26	-1.44	0.22	0.24	(0.17,0.36)
26-38	-1.73	0.21	0.18	(0.12,0.27)
>38+	-1.71	0.20	0.18	(0.12,0.20)
Maternal age				
<20 (Ref)				
20-29	1.23	0.25	3.42	(2.10,5.58)
30-39	1.13	0.29	3.11	(1.75,5.47)
40-49	0.78	0.32	2.17	(1.16,4.10)
Occupation				
Working	0.09	0.31	1.09	(0.59,2.01)
Education				
No education (Ref)				
Incomplete Primary	-0.03	0.32	0.97	(0.52,1.82)
Complete Primary	-0.52	0.36	0.59	(0.29,1.21)
Secondary or Higher	-1.19	1.23	0.30	(0.03,3.39)
Age at first birth				
<20 (Ref)				
20-29	0.22	0.12	1.25	(0.98,1.58)
30-39	0.77	0.38	2.17	(1.03,4.53)
40-49	-0.10	2.39	0.91	(0.008,97.5)
Working:Primaryincomplete	-0.09	0.36	0.91	(0.45,1.86)
Working: Complete primary	0.22	0.41	1.25	(0.56,2.77)
Working:Secondary or Higher	1.11	1.28	3.04	(0.25,37.34)

4.4 HIERARCHICAL BAYESIAN ANALYSIS

Hierarchical Bayesian analysis was sought to consider the hierarchical nature of the dataset and the source of random effect due to differences in clusters and households. Due to the complexity of the dataset, this would take a long time using MCMC (Taylor & Diggle, 2014). To avoid this, Integrated Nested Laplace Approximation (INLA)(Rue et al., 2009) method was used for hierarchical Bayesian analysis.

4.4.1 Model selection

Model selection of the best fit was important to avoid overfitting by removing other variables which are irrelevant in explaining the perinatal mortality (Rue et al., 2009). The full model with all variables was run one without random effects and the other one with random effects. The function Efxplot from package ggregplot was used to plot the covariates. Taking into account that INLA has no p-values, the significance of the variable is considered by examining the overlap of 2.5% and 97.5% posterior estimates with zero (Gómez-Rubio, 2020; Schrödle et al., 2011). *Figure* 6 below shows that BMI, previous birth interval and birth order were promising.

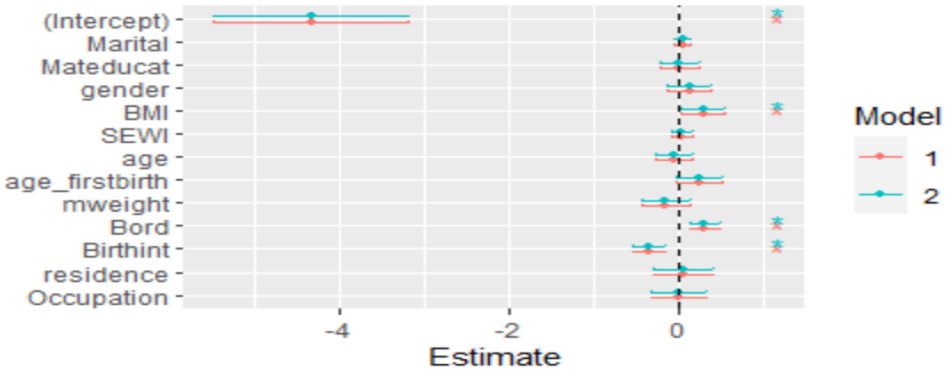


Figure 6: Significant variables for hierarchical Bayesian analysis

The full model could be reduced by removing the covariates one by one to determine the covariates with little impact by observing how each one

changes the model fit using the Deviance Information Criteria (DIC) or Akaike Information Criteria (AIC). Therefore, taking into consideration the hierarchical nature of the dataset and the source of random effect, only BMI of a woman, previous birth interval and birth order of a child were significantly associated with perinatal mortality.

1.4.2 Hierarchical Bayesian regression analysis

4.4.2.1 Hierarchical Bayesian analysis using non-informative prior

Hierarchical Bayesian regression analysis revealed birth order and BMI were positively associated with perinatal mortality. The results show a clear trend regarding body mass index and birth order favouring perinatal mortality. Perinatal mortality is more likely in overweight and obese women and the risk also increases in higher (2+) birth orders. The trend also shows that the risk of perinatal mortality is very low with previous birth intervals and the likelihood is minimal with higher previous birth intervals (+38 months). Results for hierarchical Bayesian analysis were summarized in table 6.

In this case, body mass index, birth order and previous birth interval were independently associated with perinatal mortality since the confidence interval (CI) for the adjusted odds ratios (AORs) do not contain 1. Body mass index 18.5-24.9 (Adjusted odds ratio (AOR) 2.22 95% CI (2.21-2.23), overweight women (BMI=25-29.9) AOR=2.86 95% CI(2.85,2.87), obese women AOR=2.86 95% CI(2.85,2.87), birth order 2-3 AOR=5.19 95% CI(5.186,5.193), birth order 4-5 AOR=3.39 95% CI(3.386,3.394) and birth order 6+ AOR=5.43 95% CI(5.429,5.435) were all associated with perinatal mortality. On the other hand, as it was observed in classical logistic regression previous birth interval was associated with a lower risk of perinatal mortality with AORs 0.34 95% CI(0.337,0.343), 0.34 95%

CI(0.337,0.343), 0.10 95% CI(0.08,0.118) for pregnancy intervals 15-26 months, 26-38 months and 39+ months respectively.

Table 3: INLA output for hierarchical Bayesian analysis with non-informative priors.

Variable	Mean	sd	2.5% quant	50% quant	97.5 % quant	OR	OR 95%CI
Intercept	-4.70	0.39	-5.53	-4.68	-3.99	0.009	(0.0089,0.0092)
Body mass index							
<18.5 (Ref)							
18.5-24.9	0.80	0.38	0.12	0.78	1.59	2.22	(2.20,2.24)
25-29.9	1.05	0.39	0.33	1.03	1.88	2.86	(2.81,2.91)
>30	1.05	0.44	0.23	1.04	1.95	2.86	(2.83,2.89)
Birth order							
1 (Ref)							
2-3	1.65	0.20	1.27	1.64	2.04	5.19	(5.17,5.21)
4-5	1.22	0.23	0.77	1.22	1.67	3.39	(3.37,3.40)
6+	1.69	0.27	1.16	1.70	2.21	5.43	(5.40,5.46)
Birth interval (Month)							
<15 (Ref)							
15-26	-1.09	0.18	-1.45	-1.09	-0.74	0.34	(0.335,0.338)
26-38	-1.080	0.20	-1.48	-1.08	-0.70	0.34	(0.338,0.341)
>38+	-2.31	1.10	-4.78	-2.19	-0.48	0.10	(0.096,0.102)

Furthermore, the odds of perinatal mortality happening were 2.22 times more likely in women with normal weight compared to underweight women. Also, it was 2.86 times more likely in heavyweight and obese women compared to underweight women. It was noted that perinatal mortality was 5.19 times more likely in children of order 2-3 compared to the first child, 3.39 times more likely in order 4-5 as compared to the first child and 5.43 times more likely in birth order more than six (6+) compared to first birth. In general, the likelihood of perinatal mortality increased in higher birth orders.

Analysis of previous birth interval indicated that perinatal mortality was 66.4% less likely in children born between 15-26 months compared to those born <15 months, 66% less likely in children born between 27-38 months as compared to <15 months and 90% less likely in children born >39 months as compared to <15 months. The trend shows that child spacing more than 15 months was more protective to perinatal mortality.

4.4.2.2 Hierarchical Bayesian analysis using beta conjugate priors

The results for the hierarchical Bayesian analysis by employing informative beta conjugate priors still indicates the three variables being statistically associated with perinatal mortality. With the informative priors, it was observed that the odds of perinatal mortality happening was 26 times more likely in the absence of these variables. Also, the prevalence of perinatal mortality was taking an opposite trend compared to that observed hierarchical Bayesian analysis with non-informative priors. For example, mortality was 3.53 times more likely in normal-weight women (BMI=18.5-24.9) while it was only 35% more likely in obese women (BMI=>30) with non-informative priors. A similar trend was observed in the other two variables as indicated in table 4.

Table 4: INLA output for hierarchical Bayesian analysis with informative beta priors

Variable	Mean	sd	2.5% quant	50% quant	97.5 % quant	OR	OR 95%CI
Intercept	3.27	0.67	1.99	3.26	4.63	26	(25.99,26.01)
Body mass index							
<18.5 (Ref)							
18.5-24.9	1.29	0.75	-0.14	1.27	2.81	3.53	(3.52,3.54)
25-29.9	0.79	0.80	-0.72	0.76	2.43	2.20	(2.19,2.21)
>30	0.30	0.90	-1.39	0.28	2.13	1.35	(1.34,1.37)
Birth order							
1 (Ref)							
2-3	0.90	0.78	-0.57	0.87	2.50	2.46	(2.45,2.47)
4-5	0.59	0.83	-0.96	0.56	2.27	1.80	(1.79,1.81)
6+	0.35	0.88	-1.30	0.33	2.15	1.42	(1.41,1.44)
Birth interval (Month)							
<15 (Ref)							
15-26	0.54	0.84	-1.04	0.51	2.26	1.72	(1.71,1.73)
26-38	0.45	0.86	-1.15	0.43	2.21	1.57	(1.56,1.58)
>38+	0.02	0.99	-1.91	0.02	1.97	1.02	(1.001,1.04)

4.4.2.3 Relative risk in hierarchical model with non-informative priors

The relative risk of perinatal mortality among the selected predictor variables was computed. The selected variables were those which were fitted in hierarchical Bayesian analysis. It was noted that log-link produced probability>1. To overcome this problem, we used a modified robust loga link for log-binomial in INLA and the output was summarized in table 5:

The results from table 5 shows that body mass index adjusted risk ratio (ARR=1.14 95% CI (1.13,1.14), 1.25 95% CI (1.24,1.25), 1.15 95% CI (1.14,1.15)) and birth order (ARR=2.41 95% CI (2.41,2.42), 1.79 95% CI (1.78,1.79), 2.52 95% CI (2.52,2.53) were associated with high risk of perinatal mortality. It was also noted that the risk for perinatal mortality was higher in higher birth orders (2+) with the respective adjusted risk ratio

(ARR) of 0.88 95% CI(0.87,0.88), 0.87 95% CI(0.867,0.872) and 0.56 95% CI(0.55,0.56). The previous birth interval was associated with a low risk of perinatal mortality and the risk was observed to be minimal with a higher previous birth interval (>38+).

Table 5: INLA output for relative risk in a hierarchical model

Variable	Mean	sd	2.5% quant	50% quant	97.5% quant	RR	RR 95%CI
Intercept	-4.23	0.17	-4.57	-4.23	-3.91	0.02	(0.01, 0.02)
Body mass index							
<18.5 (Ref)							
18.5-24.9	0.13	0.16	-0.17	0.13	0.44	1.14	(1.13,1.14)
25-29.9	0.22	0.18	-0.12	0.22	0.57	1.25	(1.24,1.25)
>30	0.14	0.24	-0.34	0.14	0.59	1.15	(1.14,1.15)
Birth order							
1 (Ref)							
2-3	0.880	0.123	0.639	0.880	1.121	2.41	(2.41,2.42)
4-5	0.580	0.148	0.286	0.582	0.868	1.79	(1.78,1.79)
6+	0.926	0.189	0.544	0.930	1.287	2.52	(2.52,2.53)
Birth interval (Months)							
<15 (Ref)							
15-26	-0.13	0.14	-0.41	-0.13	0.14	0.88	(0.87,0.88)
26-38	-0.14	0.16	-0.45	-0.14	0.16	0.87	(0.87,0.87)
>39+	-0.58	0.54	-1.73	-0.55	0.38	0.56	(0.55,0.56)

The relative risk of perinatal mortality among women with normal weight was 14% more likely as compared to underweight women, 25% more likely in overweight women compared to underweight women and 15% more likely in obese women compared to underweight women. The relative risk of perinatal mortality was 2.41 times more likely in children of birth order 2-3 as compared to the first birth, 79% more likely in birth order 4-5 compared to first birth and 2.52 times more likely in birth order 6+ as compared to the first birth. The trend shows that the relative risk of perinatal mortality is higher in higher birth orders (2+).

On examining the relative risk (RR) of perinatal mortality concerning previous birth interval, it was observed that perinatal mortality was 12% less likely in children born between 15-26 months as compared to those born <15 months, 13% less likely in children born between 27-38 months and 44% less likely in children born >39 months as compared to <15 months. Generally, the trend shows that child spacing more than 15 months was more protective to perinatal mortality.

CHAPTER FIVE: DISCUSSION

Perinatal mortality is a major threat to child survival in Tanzania(Mboya et al., 2020; Ogbo et al., 2019). Perinatal mortality distribution among the selected variables varied across different levels of each variable. For instance, perinatal mortality was observed to be higher in urban residence compared to rural residences, higher in obese women compared to other BMI categories, higher in women with incomplete primary education, higher in women working than those women not working, higher in women who never married, higher in women with less than <20 years of age and higher in the first birth order than other subsequent births.

Accordingly, different models were fitted to the data to determine the determinants for perinatal mortality in the country. First, the univariate chi-square test was fitted to the data using the Rao & Scott test which takes into account the design induced distortion of the asymptotic distribution of the Pearson x^2 test (Chinomona & Mwambi, 2015). The result suggests that there is an association between perinatal mortality and some variables (education, age at first birth, age, social-economic wealth index (SEWI), occupation, birth interval and birth order). However, the five non-significant variables in the univariate analysis were also included in the multivariate analysis. This is based on the fact that a non-significant variable in the univariate analysis can be significant in the multivariate model(Lo et al., 1995). In the same context, Chinomona and Mwambi,2015 included univariate non-significant variables in multivariate logistic regression, which however showed little significant contribution. This may be attributed to some factors including the influence of many missing data which was also the case in this study (Chinomona & Mwambi, 2015; Mboya et al., 2020).

In normal Bayesian Logistic regression analysis, we disregarded the complex structure of the dataset. The criteria for the best model were the model with the highest Bayes factor (Morey & Rouder, 2018; Souza & Migon, 2004). The best model, in this case, included birth order, maternal age and previous pregnancy interval being statistically significant variables. The findings of Bayesian logistic regression confirmed the relationship between maternal age, birth order and previous birth interval with perinatal mortality. It was noted that maternal age and birth order were associated with a high risk of perinatal mortality while previous birth interval was associated with a low risk of perinatal mortality. The results show a decreasing trend of perinatal mortality >20 years. This may be attributed to factors such as an increase in the experience of mothers and the maturity of the reproductive systems (McDermott et al., 1996; Mutz-Dehbalaie et al., 2014; Neal et al., 2018).

Hierarchical Bayesian analysis was performed accounting for nested sources of variability between subjects which comes from different levels of hierarchy (Khan & Shaw, 2011). Model selection accounting for the multi-stage nature of the dataset was performed and came out with three significant variables including body mass index, birth order and previous pregnancy interval. The source of the random effect, in this case, was the clusters and households. It can be seen that when put into consideration the hierarchical nature of the dataset and the sources of random effect the significant variables slightly changed with additional body mass index (BMI).

The incorporation of random effect helped to avoid overestimation of standard errors of the covariates leading to wrong statistical significance (Lee & Nelder, 2001). For instance, the coefficient for birth order 2-3 in normal Bayesian logistic regression had higher disease odds of 2.19 (AOR=8.90, 95% CI (5.81,13.74)). However, this coefficient was decreased

to 1.647(AOR=5.19, 95% CI (5.17,5.21) when the random effect was included. In addition, the analysis showed concern about body mass index whereby the odds of perinatal mortality were higher in overweight and obese women. This is an alert to policymakers to educate women of childbearing age on the importance of maintaining their normal weight to reduce perinatal mortality. It was also noted that the odds of perinatal mortality were reduced for previous pregnancy intervals and birth orders after introducing random effects. It was also noted that the three variables were significant when informative priors were used. However, the prevalence of perinatal mortality varied oppositely compared to when non-informative priors were used.

The relative risk for the log-binomial model shows a higher risk of perinatal mortality with birth order and body mass index (BMI). The analysis also suggested low relative risk for perinatal mortality with birth intervals. These findings are supported by Rahman et al., 2010 who supports the view that previous birth interval reduces the risk of perinatal mortality and in turn increase child survival. Perinatal mortality has been reported to occur at least in part from birth order factors. This is in line with the study by Ogbo (Ogbo et al., 2019) where he reported perinatal mortality being associated with fourth or higher birth ranks in Tanzania. In a similar context, high stillbirth was also associated with birth order>3 in a study conducted in Pakistan(Afshan et al., 2019).

The major limitation of this study is the absence of other unmeasured confounding factors. The study findings may be affected by unmeasured confounding factors such as baby size, healthcare access, discrete geographical inequalities etc. (Gabrysch et al., 2019; Kanyangarara et al., 2018; Kiross et al., 2019b). Additionally, this study computed the prevalence of perinatal mortality in Tanzania using just a portion of the population. This approach lacks representativeness since it may not cover

all the relevant domains of the population. Therefore, to enhance national and subgroup prevalence of perinatal mortality we need to employ appropriate statistical methods.

CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS

The findings from this study alert policy decision-makers, public health experts and other stakeholders in Tanzania taking several steps to combat the problem of perinatal mortality and making sure that the problem remains at downward trajectory (Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) & ICF, 2016). Nevertheless, as stated earlier there should be increased efforts to reduce stillbirths which plays a very big role in perinatal mortality. In particular, this study has pointed the significant predictor variables such as BMI, birth order and previous birth interval. This study stresses out the government to invest more in maternal education to address issues such as child spacing and unnecessary heavyweight as pointed out in several studies (Auger et al., 2012; Mondal et al., 2009; Rahman et al., 2010; Wehby & López-Camelo, 2017). This study showed maternal education for complete primary education and secondary or higher was associated with a low risk of perinatal mortality while ages at first birth between 30-39 and 40-49 also were associated with low risk of perinatal mortality. The results in Hierarchical Bayesian regression are in line not only with country context but what has been shown in other previous studies. The authors have shown that increasing birth interval was negatively associated with perinatal mortality (Andargie et al., 2013; Becher et al, 2004; Molitoris et al., 2019; Mondal et al., 2009). On the other hand, a small number of children born from a single mother is protective to perinatal mortality (Akombi et al., 2019) while educated women are less likely to experience perinatal mortality as well as women who got their first baby at age >20 years (Auger et al., 2012; Kiross et al., 2019a; Rahman et al., 2010; Wehby & López-

Camelo, 2017).In addition to this, the government should also invest more in strengthening health systems and endless political will.

It should be noted that large scale surveys like TDHS-2015/16 often follow a hierarchical data structure since the survey is based on multistage cluster sampling. We need to consider the source of nested variability which comes from different levels of hierarchy in our analysis. It has been noted that standard Bayesian logistic regression tends to seriously bias parameter estimates in the multilevel analysis (Khan & Shaw, 2011). Therefore, the significant predictor variables for perinatal mortality in this study were considered to be BMI, birth order and previous birth interval after taking into account the source of nested variability in the data for multistage cluster sampling.

The findings of this study call for the following recommendations: This study involved only twelve predictor variables but there are much more social-economic, community and proximate factors which may influence the occurrence of perinatal mortality. Further studies should be conducted by involving more variables in the model to determine other variables which are associated with perinatal mortality. We also recommend the government invest more in women education since this has shown to be the best remedial action against perinatal mortality. Maternal education plays a pivotal role in reducing other factors associated with perinatal mortality such as reducing the number of children, observing child spacing, visiting antenatal clinics, reducing maternal weight etc. The findings of this study argue to policy-makers, decision-makers and planners to set programs that educate women of childbearing age to avoid being overweight, ensure enough birth interval and reduce the number of children born from a single mother to reduce perinatal mortality in the country.

REFERENCES

Adigun, A. B., Gajere, E. N., Oresanya, O., & Vounatsou, P. (2015). Malaria risk in Nigeria: Bayesian geostatistical modelling of 2010 malaria indicator survey data. *Malaria Journal*, *14*(1), 156. https://doi.org/10.1186/s12936-015-0683-6

Afshan, K., Narjis, G., & Qayyum, M. (2019). Risk factors and causes of stillbirths among pregnant women in Pakistan. *African Health Sciences*, *19*(1), 1507. https://doi.org/10.4314/ahs.v19i1.24

Akombi, B. J., Ghimire, P. R., & Renzaho, A. M. (2019). Factors associated with neonatal mortality in the African Great Lakes region: A pooled analysis of national surveys. *Journal of Global Health Reports*, *3*, e2019043. https://doi.org/10.29392/joghr.3.e2019043

Akombi, B. J., & Renzaho, A. M. (2019). Perinatal Mortality in Sub-Saharan Africa: A Meta-Analysis of Demographic and Health Surveys. *Annals of Global Health*, *85*(1), 106. https://doi.org/10.5334/aogh.2348

Andargie, G., Berhane, Y., Worku, A., & Kebede, Y. (2013). Predictors of perinatal mortality in rural population of Northwest Ethiopia: A prospective longitudinal study. *BMC Public Health*, *13*(1), 168. https://doi.org/10.1186/1471-2458-13-168

Andrade, A. R., & Teixeira, P. F. (2015). Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models.

Reliability Engineering & System Safety, *142*, 169–183. https://doi.org/10.1016/j.ress.2015.05.009

Anwar, J., Torvaldsen, S., Sheikh, M., & Taylor, R. (2018). Under-estimation of maternal and perinatal mortality revealed by an enhanced surveillance system: Enumerating all births and deaths in Pakistan. *BMC Public Health*, *18*(1), 428. https://doi.org/10.1186/s12889-018-5363-3

Arató, N., Dryden, I. L., & Taylor, C. C. (2006). Hierarchical Bayesian modelling of spatial age-dependent mortality. *Computational Statistics & Data Analysis*, *51*(2), 1347–1363. https://doi.org/10.1016/j.csda.2006.02.007

Auger, N., Delézire, P., Harper, S., & Platt, R. W. (2012). Maternal Education and Stillbirth: Estimating Gestational-age-specific and Cause-specific Associations. *Epidemiology*, *23*(2), 247–254. https://doi.org/10.1097/EDE.0b013e31824587bc

Becher, H., Müller, O., Jahn, A., Gbangou, A., Kynast-Wolf, G., & Kouyaté, B. (2004). Risk factors of infant and child mortality in rural Burkina Faso. *Bulletin of the World Health Organization*, *82*(4), 265–273.

Biracyaza E* and Habimana S, B. E. and H. S. (2019). Modeling the Risks Factors Associated with Infant Mortality in Rwanda from 2011 to 2015: Analysis of Rwanda Demographic and Health Survey (RDHS)

2014/2015. *Journal of Pediatrics & Child Health Care*, *Volume 4*(Issue 1-2019).

Bonamente, M. (2017). *Statistics and Analysis of Scientific Data*. Springer New York. https://doi.org/10.1007/978-1-4939-6572-4

Cardoso, A. R. P., Araújo, M. A. L., Andrade, R. F. V., Saraceni, V., Miranda, A. E., & Dourado, M. I. C. (2016). Underreporting of Congenital Syphilis as a Cause of Fetal and Infant Deaths in Northeastern Brazil. *PLOS ONE*, *11*(12), e0167255. https://doi.org/10.1371/journal.pone.0167255

Chinomona, A., & Mwambi, H. G. (2015). Estimating HIV Prevalence in Zimbabwe Using Population-Based Survey Data. *PLOS ONE*, *10*(12), e0140896. https://doi.org/10.1371/journal.pone.0140896

Conde-Agudelo, A., Belizán, J. M., Norton, M. H., & Rosas-Bermúdez, A. (2005). Effect of the Interpregnancy Interval on Perinatal Outcomes in Latin America: *Obstetrics & Gynecology*, *106*(2), 359–366. https://doi.org/10.1097/01.AOG.0000171118.79529.a3

de Graaf, J. P., Ravelli, A. C. J., de Haan, M. A. M., Steegers, E. A. P., & Bonsel, G. J. (2013). Living in deprived urban districts increases perinatal health inequalities. *The Journal of Maternal-Fetal & Neonatal Medicine*, *26*(5), 473–481. https://doi.org/10.3109/14767058.2012.735722

De Smedt, T., Simons, K., Van Nieuwenhuyse, A., & Molenberghs, G. (2015). Comparing MCMC and INLA for disease mapping with Bayesian hierarchical models. *Archives of Public Health*, *73*(S1), O2. https://doi.org/10.1186/2049-3258-73-S1-O2

Ebenezer, E., Londhe, V., Rathore, S., Benjamin, S., Ross, B., Jeyaseelan, L., & Mathews, J. (2019). Peripartum interventions resulting in reduced perinatal mortality rates, and birth asphyxia rates, over 18 years in a tertiary centre in South India: A retrospective study. *BJOG: An International Journal of Obstetrics & Gynaecology*, *126*(S4), 21–26. https://doi.org/10.1111/1471-0528.15848

ECMS. (2016). *30th European Conference on Modelling and Simulation (ECMS 2016): Regensburg, Germany, 31 May-3 June 2016* (T. Claus, F. Herrmann, M. Manitz, & O. Rose, Eds.). Curran Associates, Inc.

Frøen, J. F., Friberg, I. K., Lawn, J. E., Bhutta, Z. A., Pattinson, R. C., Allanson, E. R., Flenady, V., McClure, E. M., Franco, L., Goldenberg, R. L., Kinney, M. V., Leisher, S. H., Pitt, C., Islam, M., Khera, A., Dhaliwal, L., Aggarwal, N., Raina, N., Temmerman, M., … Leisher, S. H. (2016). Stillbirths: Progress and unfinished business. *The Lancet*, *387*(10018), 574–586. https://doi.org/10.1016/S0140-6736(15)00818-1

Gabrysch, S., Nesbitt, R. C., Schoeps, A., Hurt, L., Soremekun, S., Edmond, K., Manu, A., Lohela, T. J., Danso, S., Tomlin, K., Kirkwood, B., & Campbell, O. M. R. (2019). Does facility birth reduce maternal and perinatal mortality in Brong Ahafo, Ghana? A secondary analysis using data on 119 244 pregnancies from two cluster-randomised controlled trials. *The Lancet Global Health*, *7*(8), e1074–e1087. https://doi.org/10.1016/S2214-109X(19)30165-2

Gayawan, E., Adebayo, S. B., & Chitekwe, S. (2014). Exclusive Breastfeeding Practice in Nigeria: A Bayesian Stepwise Regression Analysis. *Maternal and Child Health Journal*, *18*(9), 2148–2157. https://doi.org/10.1007/s10995-014-1463-6

Gelman & Yu-Sung Su. (2020). *arm: Data Analysis Using Regression and Multilevel/Hierarchical Models* (R package version 1.11-2) [Computer software]. https://CRAN.R-project.org/package=arm

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (0 ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b16018

Goldstein, H. (2010). *Multilevel Statistical Models: Goldstein/Multilevel Statistical Models*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470973394

Gómez-Rubio, V. (2020). *Bayesian inference with Inla* (1st ed.). Taylor and Francis, CRC Press.

Harrell, F. E. (2001). Binary Logistic Regression. In F. E. Harrell, *Regression Modeling Strategies* (pp. 215–267). Springer New York. https://doi.org/10.1007/978-1-4757-3462-1_10

Hespanhol, L., Vallio, C. S., Costa, L. M., & Saragiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian Journal of Physical Therapy*, *23*(4), 290–301. https://doi.org/10.1016/j.bjpt.2018.12.006

Ikua A.W. (2010). *Determinants of perinatal mortality in Kenya* [Thesis, University of Nairobi]. http://erepository.uonbi.ac.ke/bitstream/handle/11295/65501/IKUA%20_Determinants%20of%20perinatal%20mortality%20in%20Kenya?sequence=2&isAllowed=y

Jacobsson, B., Ladfors, L., & Milsom, I. (2004). Advanced Maternal Age and Adverse Perinatal Outcome: *Obstetrics & Gynecology*, *104*(4), 727–733. https://doi.org/10.1097/01.AOG.0000140682.63746.be

Jeffrey D. Sachs. (2003, October 29). *Achieving the Millennium Development Goals: Health in the Developing World*. Second Global Consultation of the Commission on Macroeconomics and Health, Geneva.

Joanne L. Shin. (2015). *A Fully Bayesian Approach to Logistic Regression* [University of California, San Diego]. https://escholarship.org/uc/item/4306n132

Kamata, A. (2001). Item Analysis by the Hierarchical Generalized Linear Model. *Journal of Educational Measurement*, *38*(1), 79–93. https://doi.org/10.1111/j.1745-3984.2001.tb01117.x

Kanyangarara, M., Chou, V. B., Creanga, A. A., & Walker, N. (2018). Linking household and health facility surveys to assess obstetric service availability, readiness and coverage: Evidence from 17 low- and middle-income countries. *Journal of Global Health*, *8*(1), 010603. https://doi.org/10.7189/jogh.08.010603

Khan, M., & Shaw, J. (2011). Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data. *J Data Sci.*, *9*. https://doi.org/10.6339/JDS.201101_09(1).0008

Kidanto, H. L., Massawe, S. N., Nystrom, L., & Lindmark, G. (2006). Analysis of perinatal mortality at a teaching hospital in Dar es Salaam, Tanzania, 1999-2003. *African Journal of Reproductive Health*, *10*(2), 72–80.

Kiross, G. T., Chojenta, C., Barker, D., Tiruye, T. Y., & Loxton, D. (2019a). The effect of maternal education on infant mortality in Ethiopia: A systematic review and meta-analysis. *PloS One*, *14*(7), e0220076. https://doi.org/10.1371/journal.pone.0220076

Kiross, G. T., Chojenta, C., Barker, D., Tiruye, T. Y., & Loxton, D. (2019b). The effect of maternal education on infant mortality in

Ethiopia: A systematic review and meta-analysis. *PloS One*, *14*(7), e0220076. https://doi.org/10.1371/journal.pone.0220076

Lee, Y., & Nelder, J. A. (2001). Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika*, *88*(4), 987–1006. https://doi.org/10.1093/biomet/88.4.987

Lisonkova, S., Haslam, M. D., Dahlgren, L., Chen, I., Synnes, A. R., & Lim, K. I. (2016). Maternal morbidity and perinatal outcomes among women in rural versus urban areas. *Canadian Medical Association Journal*, *188*(17–18), E456–E465. https://doi.org/10.1503/cmaj.151382

Lo, S. K., Li, I. T., Tsou, T. S., & See, L. (1995). [Non-significant in univariate but significant in multivariate analysis: A discussion with examples]. *Changgeng Yi Xue Za Zhi*, *18*(2), 95–101.

Madigan, D. (2005). Bayesian Multinomial Logistic Regression for Author Identification. *AIP Conference Proceedings*, *803*, 509–516. https://doi.org/10.1063/1.2149832

Mbaruku, G., van Roosmalen, J., Kimondo, I., Bilango, F., & Bergström, S. (2009). Perinatal audit using the 3-delays model in western Tanzania. *International Journal of Gynecology & Obstetrics*, *106*(1), 85–88. https://doi.org/10.1016/j.ijgo.2009.04.008

Mboya, I. B., Mahande, M. J., Obure, J., & Mwambi, H. G. (2020). Predictors of perinatal death in the presence of missing data: A birth registry-based study in northern Tanzania. *PLOS ONE*, *15*(4), e0231636. https://doi.org/10.1371/journal.pone.0231636

McCullagh. P & Nelder J.A. (1989). *Generalized Linear Models* (2nd ed., Vol. 37). Chapman & Hall.

McCulloch, C. E., & Neuhaus, J. M. (2005). Generalized Linear Mixed Models. In P. Armitage & T. Colton (Eds.), *Encyclopedia of Biostatistics* (p. b2a10021). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470011815.b2a10021

McDermott, J., Steketee, R., & Wirima, J. (1996). Perinatal mortality in rural Malawi. *Bulletin of the World Health Organization*, *74*(2), 165–171.

Mejía-Guevara, I., Zuo, W., Bendavid, E., Li, N., & Tuljapurkar, S. (2019). Age distribution, trends, and forecasts of under-5 mortality in 31 sub-Saharan African countries: A modeling study. *PLOS Medicine*, *16*(3), e1002757. https://doi.org/10.1371/journal.pmed.1002757

Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC). (2015). *Health Sector Strategic Plan (HSSP IV)*.

Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health

(MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS) & ICF. (2016). *Tanzania Demographic and Health Survey and Malaria Indicator Survey (TDHS-MIS) 2015-16*. MoHCDGEC, MoH, NBS, OCGS, and ICF. https://dhsprogram.com/publications/publication-fr321-dhs-final-reports.cfm

Minsart, A.-F., Englert, Y., & Buekens, P. (2013). Naturalization of immigrants and perinatal mortality. *The European Journal of Public Health*, *23*(2), 269–274. https://doi.org/10.1093/eurpub/cks032

Mohammed, S. H., Muhammad, F., Pakzad, R., & Alizadeh, S. (2019). Socioeconomic inequality in stunting among under-5 children in Ethiopia: A decomposition analysis. *BMC Research Notes*, *12*(1), 184. https://doi.org/10.1186/s13104-019-4229-9

Molitoris, J., Barclay, K., & Kolk, M. (2019). When and Where Birth Spacing Matters for Child Survival: An International Comparison Using the DHS. *Demography*, *56*(4), 1349–1370. https://doi.org/10.1007/s13524-019-00798-y

Mondal, Md. N. I., Hossain, Md. K., & Ali, Md. K. (2009). Factors Influencing Infant and Child Mortality: A Case Study of Rajshahi District, Bangladesh. *Journal of Human Ecology*, *26*(1), 31–39. https://doi.org/10.1080/09709274.2009.11906162

Morey & Rouder. (2018). *Computation of Bayes Factors for Common Designs.* (R package version 0.9.12-4.2.) [Computer software]. https://CRAN.R-project.org/package=BayesFactor

Morris, W. K., Vesk, P. A., & McCarthy, M. A. (2013). Profiting from pilot studies: Analysing mortality using Bayesian models with informative priors. *Basic and Applied Ecology*, *14*(1), 81–89. https://doi.org/10.1016/j.baae.2012.11.003

Mosley, W. H., & Chen, L. C. (2003). An analytical framework for the study of child survival in developing countries. 1984. *Bulletin of the World Health Organization*, *81*(2), 140–145.

Musio, M., Sauleau, E.-A., & Mameli, V. (2013). Using Integrated Nested Laplace Approximations for Modelling Spatial Healthcare Utilization. In M. Grigoletto, F. Lisi, & S. Petrone (Eds.), *Complex Models and Computational Methods in Statistics* (pp. 187–201). Springer Milan. https://doi.org/10.1007/978-88-470-2871-5_15

Muttarak, R., & Dimitrova, A. (2019). Climate change and seasonal floods: Potential long-term nutritional consequences for children in Kerala, India. *BMJ Global Health*, *4*(2), e001215. https://doi.org/10.1136/bmjgh-2018-001215

Mutz-Dehbalaie, I., Scheier, M., Jerabek-Klestil, S., Brantner, C., Windbichler, G. H., Leitner, H., Egle, D., Ramoni, A., & Oberaigner, W. (2014). Perinatal Mortality and Advanced Maternal

Age. *Gynecologic and Obstetric Investigation*, *77*(1), 50–57. https://doi.org/10.1159/000357168

Nations Unies (Ed.). (2015). *The Millennium development goals report 2015*. United Nations.

Neal, S., Channon, A. A., & Chintsanya, J. (2018). The impact of young maternal age at birth on neonatal mortality: Evidence from 45 low and middle income countries. *PLOS ONE*, *13*(5), e0195731. https://doi.org/10.1371/journal.pone.0195731

Ng, E. S., Carpenter, J. R., Goldstein, H., & Rasbash, J. (2006). Estimation in generalised linear mixed models with binary outcomes by simulated maximum likelihood. *Statistical Modelling: An International Journal*, *6*(1), 23–42. https://doi.org/10.1191/1471082X06st106oa

Novignon, J., Ofori, B., Tabiri, K. G., & Pulok, M. H. (2019). Socioeconomic inequalities in maternal health care utilization in Ghana. *International Journal for Equity in Health*, *18*(1), 141. https://doi.org/10.1186/s12939-019-1043-x

Ogbo, F. A., Ezeh, O. K., Awosemo, A. O., Ifegwu, I. K., Tan, L., Jessa, E., Charwe, D., & Agho, K. E. (2019). Determinants of trends in neonatal, post-neonatal, infant, child and under-five mortalities in Tanzania from 2004 to 2016. *BMC Public Health*, *19*(1), 1243. https://doi.org/10.1186/s12889-019-7547-x

Pateras, K & Ntzoufras,I. (2013). *Bayesian Inference and Variable Selection in Normal and Binomial Regression models with Applications in Medical Research"* [National and Kapodistrian University of Athens]. http://www2.stat-athens.aueb.gr/~jbn/courses/diplomatikes/biostats/Pateras(2013).pdf

Plummer, M., Best, N., Cowles, K., & Vines, K. (2005). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*.

R Core Team (2014). (n.d.). *R: A language and environment for statistical computing. R Foundation for Statistical Computing* (3.6.3) [Computer software]. http://www.r-project.org/

Racape, J., Schoenborn, C., Sow, M., Alexander, S., & De Spiegelaere, M. (2016). Are all immigrant mothers really at risk of low birth weight and perinatal mortality? The crucial role of socio-economic status. *BMC Pregnancy and Childbirth*, *16*(1), 75. https://doi.org/10.1186/s12884-016-0860-9

Rahman, S., Bener, Khalil Salameh, & El Ansari. (2010). Socioeconomic associations of improved maternal, neonatal, and perinatal survival in Qatar. *International Journal of Women's Health*, 311. https://doi.org/10.2147/IJWH.S12426

Rerimoi, A. J., Jasseh, M., Agbla, S. C., Reniers, G., Roca, A., & Timæus, I. M. (2019). Under-five mortality in The Gambia: Comparison of the results of the first demographic and health survey with those from

existing inquiries. *PLOS ONE*, *14*(7), e0219919. https://doi.org/10.1371/journal.pone.0219919

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*(2), 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Ruth Sharon Apio, B., Mawa, R., Lawoko, S., & Nand Sharma, K. (2019). Socio-economic Inequality in Stunting among Children Aged 6-59 Months in a Ugandan Population Based Cross-sectional Study. *American Journal of Pediatrics*, *5*(3), 125. https://doi.org/10.11648/j.ajp.20190503.18

Schrödle, B., Held, L., Riebler, A., & Danuser, J. (2011). Using integrated nested Laplace approximations for the evaluation of veterinary surveillance data from Switzerland: A case-study: Evaluation of Veterinary Surveillance Data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *60*(2), 261–279. https://doi.org/10.1111/j.1467-9876.2010.00740.x

Souza, A. D. P., & Migon, H. S. (2004). Bayesian binary regression model: An application to in-hospital death after AMI prediction. *Pesquisa Operacional*, *24*(2), 253–267. https://doi.org/10.1590/S0101-74382004000200003

Taylor, B. M., & Diggle, P. J. (2014). INLA or MCMC? A tutorial and comparative evaluation for spatial prediction in log-Gaussian Cox processes. *Journal of Statistical Computation and Simulation*, *84*(10), 2266–2284. https://doi.org/10.1080/00949655.2013.788653

Teacy, W. T. L., Luck, M., Rogers, A., & Jennings, N. R. (2012). An efficient and versatile approach to trust and reputation using hierarchical Bayesian modelling. *Artificial Intelligence*, *193*, 149–185. https://doi.org/10.1016/j.artint.2012.09.001

United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*. https://sustainabledevelopment.un.org/post2015/transformingourworld/publication

United Nations (UN). (2015). *The Millennium Development Goals Report* (p. 72). United Nations. https://en.wikipedia.org/wiki/Millennium_Development_Goals

Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth Edition) [Computer software]. Springer. url = {https://www.stats.ox.ac.uk/pub/MASS4/}

Wakefield, J. (2013). *Bayesian and Frequentist Regression Methods*. Springer New York. https://doi.org/10.1007/978-1-4419-0925-1

Wehby, G. L., & López-Camelo, J. S. (2017). Maternal Education Gradients in Infant Health in Four South American Countries. *Maternal and*

Child Health Journal, *21*(11), 2122–2131. https://doi.org/10.1007/s10995-017-2327-7

WHO. (1978). *International Conference on Primary Health Care* (CF/HST/1985-034/Anx.04/07). WHO, UNICEF. https://www.unicef.org/about/history/files/Alma_Ata_conference_1978_report.pdf

Zhu, L., Gorman, D. M., & Horel, S. (2006). Hierarchical Bayesian spatial models for alcohol availability, drug "hot spots" and violent crime. *International Journal of Health Geographics*, *5*(1), 54. https://doi.org/10.1186/1476-072X-5-54

Zhu, M., & Lu, A. Y. (2004). The Counter-intuitive Non-informative Prior for the Bernoulli Family. *Journal of Statistics Education*, *12*(2), 3. https://doi.org/10.1080/10691898.2004.11910734

Zupan, J., & Åhman, E. (2006). *Neonatal and perinatal mortality: Country, regional and global estimates*. World Health Organization.