MULTILEVEL MODELLING OF FACTORS AFFECTING CHILD MORTALITY IN KENYA

by Lucas Mwambire

Submission date: 19-Nov-2020 03:49PM (UTC+0300)

Submission ID: 1451034733

File name: ithesis.pdf (769.04K)

Word count: 11255

Character count: 59479

University of Nairobi

School of Mathematics

ISSN: 2410-1397

Master Project in Mathematics

MULTILEVEL MODELLING OF FACTORS AFFECTING CHILD MORTALITY IN KENYA

Research Report in Mathematics, Number 42, 2020

Mwambire Lucas Ruwa

November 2020

Submitted to the School of Mathematics in partial fulfillment for a degree in Master of Science in Pure Mathematics

Master Project in Mathematics

University of Nairobi
November 2020

MULTILEVEL MODELLING OF FACTORS AFFECTING CHILD MORTALITY IN KENYA

Research Report in Mathematics, Number 42, 2020

Mwambire Lucas Ruwa

School of Mathematics
College of Biological and Physical sciences
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master of Science Project
Submitted to the School of Mathematics in partial fulfillment for a degree in Master of Science in Pure Mathematics

Prepared for	The Director Graduate School University of Nairobi
Monitored by	School of Mathematics

Abstract

Background: Child mortality is among the uttermost important sensitive indicator of health and overall development of a country. Despite remarkable progress which has been made to improve child survival, child mortality still remains a problem. This study was aimed to observe the influencing determinants of child mortality in Kenya.
Methods: This study used Kenya Demographic and Health Survey(KDHS) 2014 data to investigate the predictors of child mortality. Multilevel logistic regression approach has been used to examine risk factors associated with child mortality by inclusion of unobserved random effects. The model with the best fit was checked using Akaike information criterion(AIC) and Bayesian information criterion(BIC).
Results: The study found out that breastfeeding and births in the last years had a substantial influence on child mortality ($p < 0.05$). The results also showed that model I(with household and no predictors) had the best fit for child mortality due to it's lower value of Akaike information criterion(AIC) and Bayesian information criterion(BIC) obtained.
Conclusion: The study found out that there was presence of unobserved random effects at community cluster level in which the variables included in the model cannot explain. There was no significant effect in household level even after inclusion of the explanatory variables meaning that the model can be fitted without inclusion of the household cluster level.

Dissertations in Mathematics at the University of Nairobi, Kenya.
ISSN 2410-1397: Research Report in Mathematics

DISTRIBUTOR: School of Mathematics, University of Nairobi, Kenya

Declaration and Approval

I the undersigned declare that this project report is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

Signature Date

MWAMBIRE LUCAS RUWA
Reg No. I56/88287/2016

In my capacity as a supervisor of the candidate, I certify that this report has my approval for submission.

Signature Date

Dr Idah Orowe
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: orowe@uonbi.ac.ke

Dedication

This project is dedicated to my wife; Jessica Ruwa, my son; Zayden Ruwa, my father; Dr. Thomas Mwambire Dzeha, my mother; Florence Uchi Mkangi and my siblings; Lynda Njira and Hayder Mwambire.
Most importantly I dedicate this project to the Almighty God for always giving me strength, ability to understand and protection.

Contents

Acknowledgments

First, I give thanks to Almighty God for giving me strength and ability to understand, learn and complete this project.
I would like to express deep gratitude to my guide Dr. Idah Orowe for her guidance, patience, encouragement and gracious support throughout the course of my work and for her faith in me at every stage of this research.
Lastly, I also would like to express my acknowledgement to my SST discussion group who have been instrumental in making me understand more other concepts of Multilevel analysis. In particular I would like to thank Moses, Damar, Victor and Davy for their support in this project....

Mwambire Lucas Ruwa

Nairobi, 2020.

1 Introduction

Child mortality is a vital measure of children's health, their overall well-being and overall development of a state. It is one of the important components which was included in goal number three of the recently launched Sustainable Development Goals (SDGs), where a target is set by the year 2030; child deaths which are preventable to be ended by countries (Lozano et al., 2018). According to the World Health Organization, child mortality is a probability of a child born in a specific year dying before reaching five years of age at the specific age mortality rate (Organization et al., 2011).
Various studies that have been conducted have not taken into consideration the inclusion of cluster effect which can probably lead to either overestimation or underestimation of the model parameters, hence inclusion of the cluster term will lead to the reduction of level of biasness, therefore producing unbiased estimates. The characteristics of distinctive levels like household and community have an impact on the patterns of child mortality. The use of binary logistic regression to analyze variables from distinctive levels at one single ordinary level leads to bias. The choice of multilevel modeling enables us to examine the household and community level in the same analysis instead of considering one over the other, hence making it an appropriate method for such cases. Therefore, the primary objective of this study was to identify the risk factors which are associated with under-five child mortality by inclusion of unobserved random effects using the multilevel regression technique and also to provide a predictive model for child mortality.

1.1 Background of the study

Globally, remarkable progress has been made to reduce child mortality. In the years between 1990 and 2015, there was a 53 percent decline in child mortality, it still remains a problem in many parts of the World attention (Unicef et al., 2015). It is a leading indicator of child health and overall development of a nation, as it reflects the social, economic, and environmental conditions in which children live, including their healthcare (Hill, 1991). It was one of the components of United Nations human development index, and one of the most important elements in the Millennium Development Goals (MDG)- to reduce infant and child mortality by two thirds between 1990 to 2015 (UNICEF., 2007). Under-five mortality decline has improved from 1.2 percent per year between 1990 and 1995, to 3.9 percent per year between 2005 and 2012 (UNICEF et al., 2013). In spite of this substantial drop in global child mortality rate, about 6.6 million children still die every year before their fifth birthday worldwide which implies 18,000 under-five children die each day (Hill, 1991). In 2015, 5.9 million children below five years of age died Worldwide. The risk of a child dying before completing five years of age is still the highest, with African countries leading with 81 deaths per 1000 live births which is approximately seven times higher than in the Western region which recorded eleven deaths per 1000 live births(WHO, 2012). Globally, the mortality of children below age five reduced from 12.7 million in the year 1990 to 6.3 million in 2013. Higher rates of 98 deaths per 1,000 live births have been recorded in Sub-Saharan Africa (You et al., 2013), which is fifteen times the average of first world countries (UNICEF et al., 2013). In Sub-Saharan Africa, child mortality was reduced by 54 percent (Unicef et al., 2015). Countries like Tanzania and Uganda have had a decline of 70 percent and 71 percent respectively compared to Kenya which had a 52 percent decline, meaning that, while these neighbouring countries met and surpassed their Millennium Development Goals targets, this was not met in Kenya. Findings from the KDHS 2008-09 shows the reversed and declining trend of under-five mortality. A childhood decline of 52 deaths per every 1,000 live births, down from 74 deaths per 1,000 live births reported in 2008-09 KDHS and stagnated at the same rate for the proceeding years as reported in 2014 KDHS. According to KDHS 2014, childhood mortality has declined over the years between 2003 and 2014, recording fifty two deaths and thirty nine deaths per 1,000 live births for child and infant respectively(KDHS, 2014). The records of these deaths in KDHS 2003 showed that atleast 2 in 20 children died before reaching their fifth birthday. While in 2014, it was recorded 1 child out of 20 die before hitting their fifth age, which is less than half of what was published in the 2003 KDHS. The first 28 days of life are the most at risk time for child's survival known as neonatal period. In 2013, about 44 percent of child deaths happened during this period which was up from 37 percent in 1990 (Organization et al., 2015).

Children are more at risk to many kinds of hazards than adults due to their dependency on parents or other care takers for survival. The probability of a newborn's chances of survival is dependent on whether he or she gets a recommended nutrition, immunization and hygienic environment (Skolnik, 2008). Children are not able to good care of themselves, therefore, calling for a special priority on child health to the international community. Children are prioritized internationally as indicated in The Millennium Development Goals (MDG), where three of the goals can be directly, and the other eight can be indirectly associated to child health (United Nations, 2006) (Skolnik, 2008). Child mortality in developing countries, accounts for relatively larger proportion of total number of deaths. In Kenya, approximately 8 out of 100 live births die before reaching their first birthday, signifying a huge loss of potential manpower (CBS, 2004). Over the recent years, Kenya has had an impressive and sustained decline in child mortality rate of 3 to 4 percent per annum from year 1965 to 1980 (Hill et al., 2001). This experience slowed but with a balanced improvement in the late 1980's after which the numbers plateaued and then begun to increase. The rate reduced below 2 percent between the year 1980 and 1990. Apparently, from the year 1990 and early 2000s, the declining mortality rates saw a reversal where the rates were rising or almost stagnating.

This study was build upon other various studies on child mortality which were been carried out, taking the conclusions found in those previous studies and examining factors contributing to the rate of under-five mortality in a country. Child health, and child mortality in particular, is not only an important issue in itself, but is commonly regarded as an indicator of the overall health status in a region or a country (Avogo and Agadjanian, 2010).

1.2 Statement of the problem

Childhood mortality can be prevented. The cause of death of children may be as a result of conditions and diseases that can be associated to the quality of care during the time of pregnancy, birth and growth, which can be prevented and treated (Unicef et al., 2015). Despite the drop of numbers due to the various health measures, policies, strategies and interventions put in place, child mortality in Kenya remains among the highest in the world, leading to one of the challenges the state needs to address. The intention of this undertaken study is to use the Kenya Demographic and Health Survey 2014 data to ascertain the important factors of child mortality.

Realization from the previous studies which have explored child survival in-depth is that, these studies have relied in considering all experimental units as independent, therefore any factor that affects child survivability will have similar effect in all communities or households. By applying the multilevel approach we were able to determine the effects of these factors and their variability across households or communities, therefore correcting the bias in the parameter estimates resulting from the clustered data.

1.3 Objective of the study

1.3.1 General Objective

The overall objective of this study is to model factors affecting child mortality using the Kenya Demographic and Health Survey (KDHS) data set for the year 2014.

1.3.2 Specific Objectives

i) To determine predictors of child mortality using binary logistic regression model.

ii) To examine the multi-level effects on factors affecting child mortality in community and household level of clusters.

iii) To compare child mortality in community and household level clusters using multilevel regression model.

1.4 Conceptual Framework

The following is a conceptual framework of socio-economic and demographic factors of child mortality.

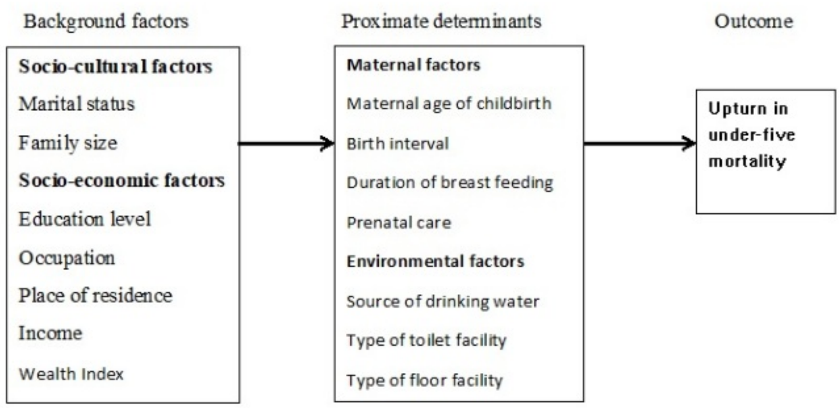


Figure 1. A conceptual framework of the study

1.5 Significance of the study

This study is expected to identify critical risk factors which are associated with child mortality, where its findings will inform and strengthen the pertinent state policies and interventions strategies which are aimed at reducing under five child mortality in Kenya.

2 Literature review

2.1 Introduction

In 1980, child survival was launched by the United Nations Children Fund (UNICEF) and World Health Organization (WHO) after the World economic meltdown. The Millennium Development Goal 4 (MDG 4) which was in relation with child survival was the focus of the development community which aimed at reducing the global rate of child mortality by two thirds between 1990 and 2015. Child mortality was included in the currently launched Sustainable Development Goals(SDGs), where it is targeted that by the year 2030, all countries to have end the preventable child deaths.

Globally, tremendous progress has been made on the reduction of under-five mortality. However, this has been unequally distributed. At the regional level, the decline in under-five mortality rates between 1990 and 2012 were over 60 percent for three WHO regions: the Americas, European and the Western Pacific regions. Mortality rates among children under the age of five remain strikingly high throughout the majority of Sub- Saharan Africa. This means that the WHO African region has increasing share of under-five deaths (WHO, 2012). There is a possibility that the mortality numbers might stagnate or even increase if no much progress will take place in the African region. Despite Sub-Saharan Africa's relatively high rates of under-five mortality, there are signs of progress in the region.

2.2 Theoretical Background

There is general consensus in the literature that a household's socio-economic and environmental characteristics do have significant effects on under-five mortality (Kembo and Van Ginneken, 2009). The survival of infants at birth and the subsequent months depends in a number of socioeconomic, demographic and biological factors. In the half of 20th century, the debate focused on which extent the socio-economic and technological changes contribute to improve the health status and mortality in developing countries (Menken, 1987).

Studies on urban-rural mortality differentials in Sub-Saharan Africa show that overall mortality, and infant and child mortality in particular, is generally lower in urban than in rural areas (Akoto and Tabutin, 1989). (APUNDA, 2016)carried out a research study in Kenya on determinants of under-five mortality, which evidently showed that children from rural areas, whose mothers had no education, and were living in poor conditions were significantly associated with child mortality. The study applied the logit regression model.

From a research study which was carried out in Ethiopia on multilevel analysis of under-five child mortality variations among it's regional states, Bedane found out that the random intercept binary logistic regression model to be the best fit of the data. This study applied the use of multilevel logistic regression model (BEDANE and HAWASSA, 2012) (Abbam, 2018) carried a study in Ghana using multilevel analysis to find factors influencing child mortality using Ghana Demographic Health Survey 2013-2014. The multilevel approach was used for determining the factors of child mortality. Findings of the study showed that both the characteristics level of child and mother like birth weight, order of birth, mother's age and level of educational were responsible for child mortality.

The mother's education level is strongly associated to child survival. Lower mortality records are recorded among those women who are highly educated, this is as a result of the exposure to vast information on contraceptives use for births spacing, good nutrition, childhood illnesses and ways of their treatment(Kamal, 2012). According to a research study which was done in Sudan in 2009, the results showed that mother's education and under-five child mortality had a significant correlation. The study applied statistical test of independence based on chi-square(Mondal et al., 2009). In a research study conducted in Ghana in 2007, found out that infants who are born to uneducated mothers are more likely to die before reaching age one compared to those infants who are born to women with primary or higher education level, the researcher used multivariate logistic model(Goro, 2007)

Garde et al found out that drinking water sources and sanitation had a negative impact on child mortality. His findings also showed maternal age to be significantly related to child mortality, that older and younger women are likely to experiences higher mortality rates than their counterparts and this depended on their reproductive maturity (Garde and Sabina, 2010). Regional differences in child mortality are related with factors at the community level which makes these regions to distinguish from one another. The health of the individuals in a community may be positively or negatively influenced by the availability and accessibility of services and social amenities in the communities. Among the factors are: prevalence of poverty, disparities in the level of community development, population density and accessibility to maternal and child health care services, which are important for child health and lessen inequities in resources and outcome of health population across regions (Irwin et al., 2007).

2.2.1 Social Demographic factors

i) Maternal age of childbirth
(Acheampong and Avorgbedor, 2017) did a research study on factors affecting child mortality in Ghana. From this research it was presumed that the decline in under-five mortality that was witnessed from the year 1988 to 2014 was not statistically

significant and that more measures needed to be put in place for further reduction. Type of assistance at delivery, maternal age and duration of breastfeeding were the strongest variables under-five mortality relative to other predictor variables. Other issues which were identified to be contributors of under-five mortality during this study were young maternal age, unskilled or no assistance at birth and exclusive breastfeeding beyond 6 months. This study applied logistic regression to predict the association between predictor variables and the dependent outcome variable.

2.2.2 Socio-economic factors

These are factors that act through proximate determinants affecting the level of morbidity and mortality.

i) Mother's level of education
Education of the mother has regularly been used as a proxy indicator of socio-economic status in most surveys and studies. It has also been correlated with hygiene, illness treatment related to early childhood morbidity (Stalling 2004). According to Root (2001), factors like lack of water or a contaminated community environment is a challenge which educated mothers may not be able to lower the risk of exposure since it is beyond their control. Nevertheless, their knowledge and wealth status enables them to seek better health care services than their counterparts who are uneducated(Root, 2001). A research study by Bello, showed that, children who were born by mothers with a higher education level experienced lower risk of infant and child mortality as compared to children born to mothers with lower education level or non-educated (Bello and Joseph, 2014).

ii) Place of delivery
Urban areas have more advantages compared to rural areas resulting to a better child survival prospects, which has been observed more often in differentials in mortality in urban and rural residence. In the third world countries, the socio-economic status, the surrounding environment and accessibility to health care services directly affect the health conditions of children residing in the rural parts, while those living in urban areas experience lower mortality rates due to ease of accessibility to proper sanitation, water and treatment(Stallings, 2004).

iii) Birth Interval
According to the United Nations statistics, nearly 40 percent of under-five children death occur in the first month of life. Two thirds of all death in the first month occurs

in the first week and two thirds of those within the first 24 hours of life. Mutunga (2004) found that child survival was found better for those who were of birth order 2-3, birth interval more than 2 years, not outcomes of multiple births. He used data from 2003 DHS in Kenya to investigate the impact of socioeconomic and environmental variables of infant and child mortality in urban areas of Kenya. The results show that the child mortality was lower for children with birth interval of more than 2 years (Mutunga, 2004). A 2008 study by Rutstein using DHS data found that the risk of mortality rapidly decreases as the birth interval increases up to 24-29 months and then decreases more slowly with longer birth intervals, but increases again for intervals of 96 or more months. The length of preceding birth interval was the explanatory variable which was measured as the number of months between the birth of the child under study and any immediately preceding birth to the mother. The study used the bivariate and multivariate approaches (Rutstein, 2008).

2.2.3 Environmental factors

Environmental factors include availability of toilet facilities, water source and method of excreta disposal. Most environmental factors are usually associated with socio-economic status of residence Rutstein (2008).

i) Type of Toilet facility
Research has shown that children living in households with better toilet facilities are less likely to fall ill compared to those without. A study by Buttenheim (2008) in Ghana, found that chances of getting infections was significantly related to toilet facility. He further stated that children who were residing in houses with toilet facilities were fifty percent less likely to get infections than those living in areas without toilet facilities (Buttenheim, 2008).

ii) Source of water
The areas with improved sources of drinking water have lower chances of getting contaminated hence reducing the spread of diseases related to water such as cholera and infections. Fayenhun (2010) used the DHS data obtained from eight countries in Sub-Saharan African, where he found out that countries like Namibia and Lesotho recorded lower child mortality as a result of accessibility to improved sources of drinking water, unlike the other group. The study showed 71 and 55 percent of households in Namibia and Lesotho respectively can access to improved sources of drinking water unlike other countries with high rate of child mortality where 40 percent or more children reside in areas where the source of drinking water was exposed to contamination(Fayehun, 2010).

3 Methodology

3.1 Introduction

This chapter focuses on the data source used for analysis, variable selection based on literature and data analysis using Multilevel approach.

3.2 Data Source

The data used in the empirical analysis was obtained from the Kenya Demographic and Health Surveys (KDHS 2014). KDHS provides information on fertility, mortality, health issues, socio-economic and environmental conditions. KDHS is a nationally representative sample of women aged 15 to 49 and men aged 15 to 54 selected from 400 clusters throughout the eight provinces in Kenya.

3.3 Data Availability

The DHS data used for this study is openly available and can be downloaded from http://www.measuredhs.com/data/available-datasets.cfm?inputSearch=KENYA.

3.4 Study Variables

3.4.1 Dependent Variable

The dependent variable of the study is the time between birth and death of a child under age five years.

3.4.2 Explanatory (Covariates) Variables

This study used variables available in the KDHS 2014 data. This included socio-economic, demographic and environmental factors which were expected to have impacts on child mortality in Kenya.

3.5 Variable Selection

The study considered factors based on extensive literature review. The essential and proximate determinants of child mortality as discussed in literature review are socio-economic, demographic and environmental factors.

3.6 Binomial logistic regression

It is a logistic regression which is typically used when the outcome variable is dichotomous (binary) and the predictors are either categorical or continuous. It is also used to obtain odds ratio in the presence of more than one explanatory variable. In this study, binomial logistic model is used to achieve one of the specific objectives which is determining the effects of predictors on the response variable (child mortality). The model is very handy when it comes to measure associations and predict outcomes, hence, it will be used to explore strong predictors influencing child mortality without the inclusion of the cluster terms (household and community level).

3.6.1 Major assumptions of binomial logistic regression

i) The response variable should be dichotomous in nature.

ii) The data should not have outliers.

iii) There should be no multicollinearity among the predictors.

iv) The sample size should be adequate.

3.6.2 Logistics Regression model

Lets consider a random variable that can take on one of two possible outcomes. With a data set of total sample size n, which has each independent observations, then Z which is a random variable is considered as a column vector of n which has variables (Z_i) which are random and binomial to Z.
Let N stand for the total population size and n taken as a column vector which constitutes n_i units representing the number of observations in population i where $i = 1$ to n. Therefore,

$$\sum_{i=1}^{n} = N$$

Let Y to be a column matrix of size N where individual unit Y_i is a random variable showing the number of success of Z for population i.We let column matrix y to contain components y_i showing the summed up number of successes observed for each population.
We let ρ to be the column matrix of length N having components $\rho_i = P(Z_i = 1|i)$, which is the probability of success in the i^{th} population for any given observation.
The linear element of the model is comprised of a design matrix and vector of parameters which are to be estimated. The design matrix of the explanatory variables X, has $K + 1$ number of columns and N number of rows, where K represent the number of predictors that are specified in the model.

In the design matrix, each of it's row has the first component x_{io} which is equal to $1(x_{io} = 1)$,which is also the intercept or α. The parameter vector, β, is regarded as a column vector of length $K+1$. One parameter is matched with each of the K columns of predictors in $X+1$, β_o, for the intercept. The model of logistic regression equates the transformation logit, the probability of success of the log-odds to the linear component.

$$log\left(\frac{\rho_i}{1-\rho_i}\right) = \sum_{k=0}^{K} x_{ik}\beta_k \qquad \textbf{(eqn 3.1)}$$

where

$$i = 1,2,3,...,N$$

3.6.3 Estimation of parameter

Binomial logistic regression was performed to assess the $K+1$ parameters(β) which are unknown in equation 3.1. The maximum likelihood estimation (MLE)will be performed to get the set of parameters for which the likelihood of the data that is observed is greatest. The equation of the maximum likelihood estimate is obtained from the probability distribution of the outcome variable.
Since every y_i constitutes a count that is binomial in the i^{th} population, the joint probability density function(pdf) of Y is shown as:

$$f(y|\beta) = \prod_{i=1}^{N}\left(\frac{n_i!}{y_i!(n_i-y_i)!}\right)\rho_i^{y_i}(1-\rho_i)^{(n_i-y_i)} \qquad \textbf{(eqn 3.2)}$$

The chances of a success for each of the n_i trials is ρ_i, and probability of y_i successes is $\rho_i^{y_i}$. Similarly, the probability of failures $n_i - y_i$ is $(1-\rho_i)^{n_i-y_i}$. The pdf in equation 3.2 express the y values as a function of β values which are fixed.
The likelihood function has a similar format as the probability density function(pdf), apart from the reversed parameters β in terms of known y values which are also fixed. Therefore,

$$L(\beta|y) = \prod_{i=1}^{N}\left(\frac{n_i!}{y_i!(n_i-y_i)!}\right)\rho_i^{y_i}(1-\rho_i)^{(n_i-y_i)} \qquad \textbf{(eqn 3.3)}$$

The estimates of maximum likelihood are the figures for β that maximizes the likelihood in equation 3.3. To compute for the maximum likelihood estimates we need to find the 1^{st} and 2^{nd} derivatives of the likelihood function.
Simplifying the likelihood equation 3.3, where ρ_i are constants to be neglected. Since $a^{x-y} = \frac{a^x}{a^y}$.
Therefore, the equation to the maximized, that is after terms have been rearranged can be written as:

$$\prod_{i=1}^{N}\left(\frac{\rho_i}{1-\rho_i}\right)^{y_i}(1-\rho_i)^{n_i} \qquad \textbf{(eqn 3.4)}$$

Taking the exponential (e) on each sides of equation, we get

$$\left(\frac{\rho_i}{1-\rho_i}\right) = e^{\sum_{k=0}^{K} x_{ik}\beta_k} \qquad \textbf{(eqn 3.5)}$$

Working out for ρ_i becomes,

$$\rho_i = \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right) \qquad \textbf{(eqn 3.6)}$$

Substituting equation 3.5 as the 1^{st} term and equation 3.6 as the 2^{nd} term, equation 3.4 then becomes,

$$\prod_{i=1}^{N}\left(e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right)^{y_i}\left(1-\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right)^{n_i} \qquad \textbf{(eqn 3.7)}$$

$(a^x)^y = a^{xy}$ will be used to simplify the 1^{st} product and replace the value 1 with $\frac{1+e^{\sum x\beta}}{1+e^{\sum x\beta}}$ for the 2^{nd} product to be simplified. Now equation 3.7 can be written as:

$$\prod_{i=1}^{N}\left(e^{y_i\sum_{k=0}^{K} x_{ik}\beta_k}\right)\left(1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right)^{-n_i} \qquad \textbf{(eqn 3.8)}$$

Equation 3.8 is the core likelihood function for maximizing. Therefore, taking the natural log of the above equation 3.8 gives log likelihood function:

$$l(\beta) = \sum_{i=1}^{N} y_i\left(\sum_{k=0}^{K} x_{ik}\beta_k\right) - n_i \cdot log\left(1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right) \qquad \textbf{(eqn 3.9)}$$

We will put the 1^{st} derivative with respect to every β equal to zero to get the critical points of the log likelihood function. Thus, performing differentiation in equation 3.9. That,

$$\frac{\partial}{\partial\beta_k}\sum_{k=0}^{K} x_{ik}\beta_k = x_{ik} \qquad \textbf{(eqn 3.10)}$$

We will use the general rule $\frac{\partial}{\partial x} logy = \frac{1}{y}\frac{\partial y}{\partial x}$ in differentiating the second half of equation 3.9. Therefore, differentiating equation 3.9 with respect to every β_k, we get

$$\begin{aligned}
\frac{\partial l(\beta)}{\partial \beta_k} &= \sum_{i=1}^{N} y_i x_{ik} - n_i \cdot \frac{1}{1+e^{\sum_{i=1}^{K} x_{ik}\beta_k}} \cdot \frac{\partial}{\partial \beta_k}\left(1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right) \\
&= \sum_{i=1}^{N} y_i x_{ik} - n_i \cdot \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} \cdot e^{\sum_{k=0}^{K} x_{ik}\beta_k} \cdot \frac{\partial}{\partial \beta_k} \sum_{k=0}^{K} x_{ik}\beta_k \\
&= \sum_{i=1}^{N} y_i x_{ik} - n_i \cdot \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} \cdot e^{\sum_{k=0}^{K} x_{ik}\beta_k} \cdot x_{ik} \\
&= \sum_{i=1}^{N} y_i x_{ik} - n_i \rho_i x_{ik}
\end{aligned} \qquad \textbf{(eqn 3.11)}$$

The MLE for β will be obtained through putting individual $K+1$ equations in equation 3.11 to be equal to zero and finding for each β_k. The matrix will be formed through differentiating each $K+1$ equations in equation 3.11 for a second time with respect to each parameter β indicated by β_k.

The matrix of the 2nd derivative takes the general form:

$$\begin{aligned}
\frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_{k'}} &= \frac{\partial}{\partial \beta_{k'}} \sum_{i=1}^{N} y_i x_{ik} - n_i x_{ik} \rho_i \\
&= \frac{\partial}{\partial \beta_{k'}} \sum_{i=1}^{N} -n_i x_{ik} \rho_i \\
&= -\sum_{i=1}^{N} n_i x_{ik} \frac{\partial}{\partial \beta_{k'}} \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} \right)
\end{aligned} \qquad \textbf{(eqn 3.12)}$$

To solve equation 3.12 we will come up with two general differentiation rules. The 1^{st} rule for differentiation of exponential functions:

$$\frac{d}{dx} e^{u(x)} = e^{u(x)} \frac{d}{dx} u(x) \qquad \textbf{(eqn 3.13)}$$

We let $u(x) = \sum_{k=0}^{K} x_{ik}\beta_k$, in this case. The 2^{nd} rule, is the quotient rule for differentiation of exponential functions:

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)\cdot f'(a) - f(a)\cdot g'(a)}{[g(a)]^2} \qquad \textbf{(eqn 3.14)}$$

The application of these two rules will enable us to solve equation 3.12

$$\begin{aligned}
\frac{d}{dx}\frac{e^{u}(x)}{1+e^{u(x)}} &= \frac{1+e^{u(x)}\cdot e^{u(x)}\frac{d}{dx}u(x)-e^{u(x)}\cdot e^{u(x)}\frac{d}{dx}u(x)}{(1+e^{u(x)})^2} \\
&= \frac{e^{u(x)}\frac{d}{dx}u(x)}{(1+e^{u(x)})^2} \\
&= \frac{e^{u(x)}}{1+e^{u(x)}}\cdot\frac{1}{1+e^{u(x)}}\cdot\frac{d}{dx}u(x)
\end{aligned} \quad \textbf{(eqn 3.15)}$$

Therefore, the equation 3.12 can be expressed as:

$$\frac{\partial^2 l(\beta)}{\partial\beta_k\partial\beta_{k'}} = -\sum_{i=1}^{N} n_i x_{ik}\rho_i(1-\rho_i)x_{ik'} \quad \textbf{(eqn 3.16)}$$

3.7 Multilevel Logistic Regression Model

In this study, a multilevel model was adopted to model child mortality variations among the regions of Kenya, where it was used to achieve our second objective of determining the random effects on factors associated with child mortality both at household and community level. Multilevel models are also known as hierarchical linear models, linear mixed models or mixed-effect models. This model will be adopted to establish whether the cluster term has a significant effect on the model and also to be used under the circumstances where the inclusion of a cluster term will be necessary to produce unbiased estimates.

Four models will be fitted to analyze the effect of the cluster term on the response variable, with and without inclusion of the predictors. The first model (empty model 1) to be estimated with the household level without inclusion of explanatory variables. The second model (empty model 2) will consider only the community level without inclusion of the predictors to determine its effect on the response variable (child mortality). The third model (model 3) will incorporate the household level with inclusion of the predictors to measure the cluster effect on the response variable with added predictors. And the fourth model (model 4) will be fitted by including the predictors with inclusion of the community level. The fixed and random effects in this study are the paramount concepts in multilevel analysis and will therefore be applied in results interpretation.

3.7.1 Modification for cluster effect

Let ρ_{ij} be the probability that child i living in region j died before reaching age five (binary response), which depend on the individual level predictor variable X_{ij} and Z_j which is the group-level predictor variable. Therefore, we will introduce u_{0j} and u_{ij} as variables into the model to adjust for cluster effect. Where u_{0j} is the level-1 residual and u_{1j} is the cluster effect(level-two). The cluster effect is considered to have a normal distribution with mean zero and a variance which is constant.

Empty multilevel logistic model

The empty model (intercept model) is expressed as:

$$logit\left(\rho_{ij}\right) = \beta_{00} + u_{0j} \qquad \textbf{(eqn 3.17)}$$

where β_{00} is the average intercept of the outcome variable, u_{0j} is the random part of the model (level-2 residual) which is taken to be mutually independent with a normal distribution of mean zero and variance σ_0^2, and i-Indicate level-one unit (individual), j-indicates level two unit (region).

Full model

The full model is specified as

$$logit\left(\rho_{ij}\right) = \beta_{00} + \beta_{01}Z_j + \beta_{10}X_{ij} + u_{0j} + u_{1j}X_{ij} \qquad \textbf{(eqn 3.18)}$$

Where, $u_{1j} \sim N(0, \sigma_u^2)$. The intercept β_{00} and the regression coefficients β_{01} and β_{10} are fixed effects, whereas u_{0j} and u_{1j} are the random effects of level Z, X_{ij} is the explanatory variable at level one and $u_{0j} + u_{1j}X_{ij}$ are the random part of the model that is considered as the interaction between the group(region) and explanatory variables.

3.7.2 Intra-class Correlation Coefficient (ICC)

As we are doing a multilevel modeling, it is therefore necessary for us to compute an Intra-class correlation coefficient (ICC) to enable us to evaluate whether the variation in the outcome is essentially within or between clusters. The ICC is calculated as a proportion of error variance in the group level over the total error variance. The value obtained after calculating the ICC will determine the level of clustering in the data, therefore, any value obtained that is greater than 0.05 (ICC>0.05) meaning the cluster variable has a significant effect on response variable, meaning that clustering is taking place. The larger the value after calculating for ICC, the more indictive of potential clustering taking place which can be flushed out by inclusion of predictors to see how much variability can be accounted for.

The intra-class correlation coefficient is:

$$\rho = \frac{\sigma_{uo}^2}{\sigma_{uo}^2 + \sigma_e^2} \qquad \textbf{(eqn 3.19)}$$

Where σ_{uo}^2 is the intercept variance (level-2), σ_e^2 is the variance of the residual and ρ is the Intra-class Correlation Coefficient. Therefore, the Intra-class correlation coefficient (ICC) is used to show the amount of variability which is not explained by the any of the predictors in the model that can be accredited to the cluster variable as contrasted to the overall variance which is unexplained.

3.8 Goodness of Fit of the models

To determine the most appropriate model with the best fit Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics were employed. These two information criteria work in the same way but impose different penalties when it comes to the introduction of too many variables.

Akaike Information Criterion (AIC)

The AIC can be expressed as;

$$AIC = -2(Log - likelihood) + 2k$$

Where k represents the number of model parameters used to build the model including the intercept and the log-likelihood is a measure of model fit, in that, a higher number indicates a better fit which is got from the statistical output.

A smaller value of Akaike Information Criterion (AIC) indicate a good fit (Boco, 2010) and that the AIC value will always increase if an unwarranted variable has been added in the model.

Bayesian Information Criterion (BIC)

It is also known as Schwarz Criterion and is used for model selection with the best fit of data. The BIC solves the problem of overfitting by presenting a penalty term for the number of parameters in the model. This problem comes as a result of adding parameters when fitting models.
The BIC can be expressed as follows;

$$BIC = -2LL + klog(n)$$

Where LL-is the log-likelihood of the model, k is the number of model parameters, and n is the sample size.
When it comes to choosing the model, the one with the lowest BIC value after calculating for each model is referred as the best model.

3.9 Data analysis

Multilevel logistic regression was used to determine the likelihood of the predictor variables affecting under-five mortality in the presence of cluster terms. A 5% level of statistical significance will be used throughout this study. The data analysis will be done using statistical (software) packages namely R and STATA. The R software will be used to conduct all the statistical analysis (Multilevel analysis) while STATA will be used for coding and creating dummy variables.

4 Results

Table 1. Demographic distribution of child survivability

	Child Survivability		
Explanatory Variables	**Alive (n=20,093)**	**Dead (n=871)**	**Total**
Age at First Birth			
6-15	2259(95.52%)	106(4.48%)	2365(100%)
16-25	16680(95.88%)	717(4.12%)	17397(100%)
26-35	1128(95.92%)	48(4.08%)	1176(100%)
36-45	26(100%)	0	26(100%)
Region			
Coast	2531(95.51%)	119(4.49%)	2650(100%)
North Eastern	1538(96.49%)	56(3.51%)	1594(100%)
Eastern	2906(96.38%)	109(3.62%)	3015(100%)
Central	1356(95.49%)	64(4.51%)	1420(100%)
Rift valley	6618(96.61%)	232(3.39%)	6850(100%)
Western	1889(95.55%)	88(4.25%)	1977(100%)
Nyanza	2757(94.22%)	169(5.78%)	2926(100%)
Nairobi	498(93.61%)	34(6.39%)	532(100%)
Place of residence			
Urban	6532(95.66%)	296(4.34%)	6828(100%)
Rural	13561(95.93%)	575(4.07%)	14136(100%)

Table 2. Demographic distribution of child survivability

	Child Survivability		
Explanatory Variables	**Alive (n=20,093)**	**Dead (n=871)**	**Total**
Source of drinking water			
Piped	6686(95.84%)	290(4.16%)	6976(100%)
Borehole	7517(96.04%)	310(3.96%)	7827(100%)
River	5174(95.64%)	236(4.36%)	5410(100%)
Other	716(95.34%)	35(4.66%)	751(100%)
Toilet facility			
Flush toilet	1584(96.23%)	62(3.77%)	1646(100%)
Pit latrine	13055(95.75%)	579(4.25%)	13634(100%)
No facility	5454(95.95%)	230(4.05%)	5684(100%)
Educational			
No education	4406(96.1%)	179(3.9%)	4585(100%)
Primary	10551(95.44%)	504(4.56%)	11055(100%)
Secondary	3857(96.35%)	146(3.65%)	4003(100%)
Tertiary	1279(96.82%)	42(3.18%)	1321(100%)
Wealth Index			
Poor	11047(95.84%)	479(4.16%)	11526(100%)
Middle	3334(95.34%)	163(4.66%)	3497(100%)
Rich	5712(96.15%)	229(3.85%)	5941(100%)
Births in last five years			
1	9473(97.53%)	240(2.47%)	9713(100%)
2	8568(95.37%)	416(4.63%)	8984(100%)
3	1943(90.96%)	193(9.04%)	2136(100%)
4	100(86.21%)	16(13.8%)	116(100%)
5	9(60%)	6(40%)	15(100%)
Breastfeeding Status			
No	14570(95.21%)	733(4.79%)	15303(100%)
Yes	5523(97.56%)	138(2.44%)	5661(100%)

Table 3. Demographic distribution of child survivability

	Child Survivability		
Explanatory Variables	**Alive (n=20,093)**	**Dead (n=871)**	**Total**
Marital status			
Single	1268(97.46%)	33(2.54%)	1301(100%)
Married	17042(95.9%)	728(4.1%)	17770(100%)
Widowed	461(92.76%)	36(7.24%)	497(100%)
Divorced	1322(94.7%)	74(5.3%)	1396(100%)
Work status			
No	8009(95.78%)	353(4.22%)	8362(100%)
Yes	12084(95.89%)	518(4.11%)	12602(100%)
Birth order			
1-2	8839(96.16%)	353(3.84%)	9192(100%)
3-4	5908(96.14%)	237(3.86%)	6145(100%)
5+	5346(95.01%)	281(4.99%)	5627(100%)
Place of delivery			
Non-facillity	10076(95.82%)	440(4.18%)	10516(100%)
Facility	10017(95.87%)	431(4.13%)	10448(100%)

Table 4. Multilevel modelling analysis of factors associated with child mortality in Kenya

	Model I		Model II		Model III		Model IV		Model V	
Explanatory Variables	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value
Intercept	0.041	0.000	0.040	0.000	-0.014	0.179	-0.013	0.215	-0.014	0.175
Age at First Birth (Ref: 6 -15)										
16-25					-0.0003	0.928	-0.0008	0.855	-0.0003	0.93
26-35					0.004	0.53	0.003	0.616	0.004	0.536
36-45					-0.04	0.299	-0.04	0.299	0.041	0.292
Region (Ref: Coast)										
North Eastern					-0.02	0.003	-0.019	0.005	-0.02	0.002
Eastern					-0.001	0.717	-0.002	0.668	-0.002	0.682
Central					0.008	0.203	0.008	0.233	0.008	0.204
Rift valley					-0.01	0.028	-0.01	0.029	-0.01	0.021
Western					0.0003	0.95	0.0004	0.945	0.0002	0.969
Nyanza					0.011	0.04	0.012	0.034	0.012	0.032
Nairobi					0.025	0.011	0.024	0.016	0.025	0.01
Place of residence (Ref:Urban)										
Rural					-0.004	0.223	-0.004	0.217	-0.004	0.222
Source of drinking water (Ref: Piped)										
Borehole					-0.004	0.22	-0.004	0.222	-0.004	0.207
River					0.000	0.988	0.0001	0.971	0.0001	0.973
Other					0.003	0.63	0.003	0.658	0.004	0.617
Toilet facility (Ref: Flush toilet)										
Pit latrine					0.005	0.363	0.005	0.394	0.005	0.358
No facility					0.002	0.742	0.002	0.757	0.002	0.757
Educational (Ref: No education)										
Primary					0.003	0.419	0.004	0.379	0.003	0.412
Secondary					-0.0002	0.97	-0.0002	0.966	-0.0001	0.976
Tertiary					-0.002	0.764	-0.002	0.752	-0.002	0.764
Wealth Index (Ref: Poor)										
Middle					0.007	0.071	0.007	0.079	0.007	0.068
Rich					0.001	0.698	0.001	0.716	0.001	0.693
Breastfeeding Status (Ref: No)										
Yes					-0.033	0.00	-0.033	0.00	-0.033	0.00

Table 5. Multilevel modelling analysis of factors associated with child mortality in Kenya

	Model I		Model II		Model III		Model IV		Model V	
Explanatory Variables	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value
Marital status(Ref: Single)										
Married					0.00	0.998	0.0009	0.874	0.00	0.989
Widowed					0.034	0.001	0.034	0.001	0.034	0.001
Divorced					0.016	0.035	0.016	0.034	0.016	0.035
Work status (Ref: No)										
Yes					-0.003	0.3	-0.002	0.328	-0.003	0.288
Birth order (Ref: 1 - 2)										
3-4					-0.005	0.099	-0.006	0.063	-0.005	0.105
5+					0.005	0.153	0.004	0.229	0.005	0.151
Place of delivery (Ref: Non-facility)										
Facility					0.001	0.616	0.001	0.672	0.001	0.595
Births in last five years					0.038	0.00	0.038	0.00	0.038	0.00

(* Significant at the p<0.05 level, ref- is reference category)

Table 6. Intraclass correlation coefficient (ICC)

	Intraclass Correlation Coefficient (ICC)			
Random effects	**Model I**	**Model II**	**Model III**	**Model IV**
Household	0.008		0.003	
Community		0.078		0.06

(ICC>0.05 implies that the cluster variable has significance on child mortality)

Table 7. Model Fitness

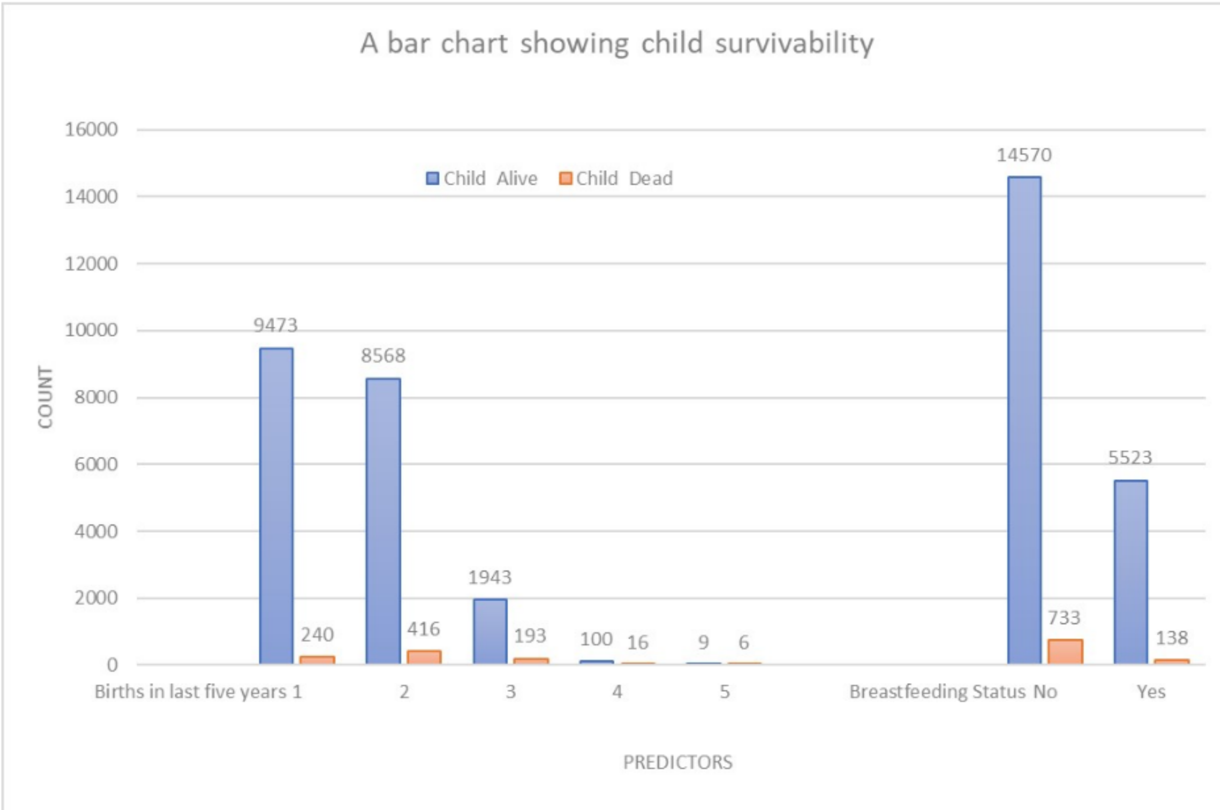
Fitness of Models	**Model I**	**Model II**	**Model III**	**Model IV**	**Model V**
AIC	-8086.765	-8167.86	-8488.919	-8540.277	-8488.7
BIC	-8062.913	-8144.008	-8226.55	-8277.909	

The Table 1, 2 and 3 shows the prevalence of child mortality across the categories of the predictors which were estimated in terms of numbers and percentages. In this study, a multilevel regression analysis was applied, where a total of five models were fitted with variables of interest which were based on literature as shown in Table 4 and table 5. The fixed effects which are measures of association between explanatory variables and child death were reported in terms of odds ratio and p-values(refer to table 4), while random effects which are a measure of variation were showed in terms of intra-class correlation coeficient(ICC) as displayed in table 6. The checking of model fitness was assessed using Akaike Information Criterion and Bayesian Information Criterion has shown in table 7.

Model I(null or empty model) had no independent variables and it is was used to test for random effects at household level clustering in the intercept and to evaluate the intra-class correlation coefficient. Model II(null or empty model) was fitted without the explanatory variables to perform for random effects at community level clustering in the intercept and to evaluate the intraclass correlation coefficient. Model III was fitted to find the effects of the independent variables on child mortality with inclusion of household cluster term. Model IV was fitted with inclusion of the explanatory variables and community level cluster term to find its effects on under-five child mortality. In Model V it contained only the independent variables to find their significance on child mortality without the inclusion of the household and community level cluster terms.

Binary logistic regression analysis

Model V: This model included only the predictors for child mortality without the inclusion of the cluster variables (household or community). The results showed in Table 4 and 5, revealed that breastfeeding status and births in last five years had a significant impact on child mortality(p-value=0.00). The log of odds of dying of children who are being breastfed compared to those who are not decreases by 0.033 controlling for other factors, and the log of odds of dying for children increases by 0.038 for each year increase in births holding other factors constant. As per the bar chart (figure 2), children deaths were higher among those who were not breastfed compared to those who were and the number of deaths children who were born in a short interval was higher compared to those who were born after some years.



Multilevel logistic regression analysis

Model I: This is an empty model which has no explanatory variables but inclusion of household cluster. The results indicated in table 6 under this model revealed that the household cluster term was found to be insignificant ($ICC < 0.05$), meaning that we don not need the household cluster to establish factors that have influence on mortality at household level.

Model II: This is also an empty model which has only included the community cluster with no predictors. As per table 4, the results shows a significant variation in child mortality at community level ($p - value < 0.05$). And the calculated ICC of this model(0.078), indicates that there is presence of clustering ($ICC > 0.05$)at community level.

Model III: All the explanatory variables with presence of household level cluster were considered in this model. The results revealed that breastfeeding status and births in last five years had an impact on child mortality($p - value < 0.05$). Keeping other predictors constant, the log of odds of dying of children who are being breastfed to those who are not reduces by 0.033 and the log of odds of dying for children increases by 0.038 for every year a child is born controlling for other factors. In this model, the household cluster could not show a significant effect to child mortality in the presence of explanatory variables($ICC < 0.05$).

Model IV: In this model all the predictors were included with presence of community cluster level. After the inclusion of the explanatory variables, the ICC was 0.06 which implied that 6% of the total variation of child mortality in Kenya can be assigned to the community level, making it to be significant indicating presence of clustering. Under this model, the results showed that births in the last five years and breastfeeding as important factors which are linked to child mortality in the presence of community cluster level($p - value = 0.00$), this means that there are unseen random factors which affects child mortality at cluster level and cannot therefore be explained by the included covariates in the model. In reference to breastfeeding status, the log of odds of dying for breastfed children compared to those who are not decreases by 0.033 holding other factors constant, while for births which took place in the last five years, the log of odds of dying for children increases by 0.038 for each year increase in births holding other factors constant.

From Model III, IV, and V; the included independent variables such as age at first birth, region, source of drinking water, place of residence, toilet facility, education level, wealth index, marital status, work status, birth order, and place of delivery were insignificant factors of child mortality. In variables like region and marital status showed some significance to child mortality in some of their categories as showed in table 4 and 5.

Model fit statistics

The model fitness test was performed using AIC and BIC to find out which of the five fitted model had the best fit. The comparison between these five fitted models is more visible when we do the analysis of AIC and BIC, since these statistical measures give a clear information on which model out of the five is the best. According to the results shown in table 7, the information got revealed that Model I had the best fit, and this is because of its calculated lower value of AIC and BIC obtained(-8086.765 and -8062.913 respectively). The AIC and BIC showed progressive increase in AIC and BIC from Model I to Model IV, meaning that the predictor value of the model reduces from Model I to Model IV.

5 Discussion

The objective of this study were to examine the factors determining child mortality with inclusion of the household and community level cluster to establish their unobserved random effects using multilevel regression model. The use of multilevel analysis helped to understand the effect of community level in child mortality. It was found that community level random intercept had an influence on the patterns of child mortality, meaning that in this model there was presence of unobserved random effects at this cluster level in which the variables included in the model cannot elucidate. This study can also be related to other studies done before by (Abera and Adjiwanou, 2017) and Adedini et al. (2015).

From the findings of the study, breastfeeding status and births in the last five years were very important in explaining child survival. The study established that children who are not being breastfed had a higher risk of dying before reaching their fifth birthday compared to their counterpart who are being breastfed. From previous studies done, breastfeeding has been found to reduce child mortality. These findings not different from those of (Bello and Joseph, 2014) where the results revealed that breastfeeding reduced child mortality in Atiba local government area in Nigeria. The study further suggested that women should be educated on the importance of breastfeeding their children especially at infancy to enhance their life since it has medically been proven. It is through breastfeeding that new born babies get their continued health benefits through the additional antibodies obtained from the breast milk enhancing their immune system. The finding are in line with WHO/UNICEF,2009 study which was showed that breast milk had nutrients, antibodies, and hormones which were very necessary for child survival and development. A research conducted in Ghana (Acheampong and Avorgbedor, 2017) showed that those children who were being breastfed had a higher chanced of reaching their fifth birthday compared to those who were not. Also the results of the research done by (Fotso and Kuate-Defo, 2005) in developing countries where they found that breastfeeding was playing a vital role in protecting infants against diseases which are infectious, which are major causes of child mortality. A study conducted in Ghana by (Edmond et al., 2006) showed that those infants who had been delayed to be breastfed were susceptible to infection related diseases leading to mortality.

Births in the last five years established to be a significant influence in child mortality. Prior studies showed that child mortality was lower for children with birth spacing of more than two years (Mutunga, 2004). A study done by Rutstein using DHS data, found out that higher chances of child mortality swiftly reduces as birth interval increases to 24-29 months and then reduces more gradually with longer birth intervals (Rutstein, 2008).

This finding is also in agreement with the results of previous studies conducted where child spacing plays a vital role when it comes to the health of the children (Rutstein, 2005) (Jahn et al., 2006) (Titaley et al., 2008), in that shorter birth intervals are related with the depletion of maternal nutrients and deficiencies.

5.1 Conclusion

This study suggested that breastfeeding status and births in the last five years were the significant determinants of child mortality in Kenya. However, factors like , place of residence, source of drinking water, toilet facility, wealth index, work status, birth order and place of delivery were found to be insignificant determinants of child mortality. It was also found that there was presence of unobserved random effects at community cluster level in which the variables included in the model cannot explain. There was no significant effect in household level even after inclusion of the explanatory variables meaning that the model can be fitted without inclusion of the household cluster level.

5.2 Future research

Further research can be done using the next KDHS data to find and understand essential determinants of child mortality to strengthen the findings of this research.

5.3 Limitations of the study

This study was based on KDHS 2014 data which is a cross-sectional study. Since secondary data was used, then there is a possibility of problems which might be caused by unobserved confounding factors. This is where there is no data availability for children of women who died, either as a result of memory lapse, misplaced data or possible data collection errors leading to missing variable values. There are also cases of under-reporting of child deaths, where the only interviewed were mothers alive. In this study there might be a case of overestimation in the observed relationship between breastfeeding and child mortality, for instance those babies whose early death may not have been attributed by not being breastfed.

References

Abbam, A. (2018). A multilevel analysis of factors influencing child mortality in ghana.

Abera, S. and Adjiwanou, V. (2017). Multilevel analysis of infant mortality and its risk factors in south africa. *International Journal of Population Studies*, 3:43.

Acheampong, G. K. and Avorgbedor, Y. E. (2017). Determinants of under five mortality in ghana; a logistic regression analysis using evidence from the demographic and health survey (1988-2014). *American Journal of Public Health*, 5(3):70–78.

Adedini, S. A., Odimegwu, C., Imasiku, E. N., Ononokpono, D. N., and Ibisomi, L. (2015). Regional variations in infant and child mortality in nigeria: a multilevel analysis. *Journal of biosocial science*, 47(2):165–187.

Akoto, E. and Tabutin, D. (1989). Socioeconomic and cultural inequalities in mortality.

APUNDA, R. (2016). *DETERMINANTS OF CHILD MORTALITY IN KENYA*. PhD thesis, School of Economics, University of Nairobi.

Avogo, W. A. and Agadjanian, V. (2010). Forced migration and child health and mortality in angola. *Social Science & Medicine*, 70(1):53–60.

BEDANE, A. S. and HAWASSA, E. (2012). Multilevel logistic modelling of under-five child mortality variations among regional states of ethiopia.

Bello, R. and Joseph, A. (2014). Determinants of child mortality in oyo state, nigeria. *African Research Review*, 8(1):252–272.

Boco, A. G. (2010). Individual and community level effects on child mortality: an analysis of 28 demographic and health surveys in sub-saharan africa. *Demographic and Health Surveys Working Paper*.

Buttenheim, A. M. (2008). The sanitation environment in urban slums: implications for child health. *Population and environment*, 30(1-2):26–47.

CBS, M. (2004). Kenya demographic and health survey 2003. central bureau of statistics (cbs)[kenya], ministry of health (moh)[kenya], and orc macro. calverton, maryland usa.

Edmond, K. M., Zandoh, C., Quigley, M. A., Amenga-Etego, S., Owusu-Agyei, S., and Kirkwood, B. R. (2006). Delayed breastfeeding initiation increases risk of neonatal mortality. *Pediatrics*, 117(3):e380–e386.

Fayehun, O. A. (2010). Household environmental health hazards and child survival in sub-saharan africa.

Fotso, J.-C. and Kuate-Defo, B. (2005). Measuring socioeconomic status in health research in developing countries: should we be focusing on households, communities or both? *Social Indicators Research*, 72(2):189–237.

Garde, M. and Sabina, N. (2010). Inequalities in child survival: looking at wealth and other socio-economic disparities in developing countries. *UK: Save the Children*.

Goro, I. (2007). The stalling child mortality in ghana: the case of the three northern regions. *University of Cape Town, Cape Town, South Africa*.

Hill, K. (1991). Approaches to the measurement of childhood mortality: a comparative review. *Population Index*, pages 368–382.

Hill, K., Bicego, G., and Mahy, M. (2001). Childhood mortality in kenya: An examination of trends and determinants in the late 1980s to mid 1990s.

Irwin, L. G., Siddiqi, A., Hertzman, G., et al. (2007). *Early child development: a powerful equalizer*. Human Early Learning Partnership (HELP) Vancouver, BC.

Jahn, A., Kynast-Wolf, G., Kouyaté, B., and Becher, H. (2006). Multiple pregnancy in rural burkina faso: frequency, survival, and use of health services. *Acta obstetricia et gynecologica Scandinavica*, 85(1):26–32.

Kamal, S. M. (2012). Maternal education as a determinant of neonatal mortality in bangladesh. *Journal of Health Management*, 14(3):269–281.

KDHS (2014). Infant and under-five mortality rate. Technical report, DHS.

Kembo, J. and Van Ginneken, J. K. (2009). Determinants of infant and child mortality in zimbabwe: Results of multivariate hazard analysis. *Demographic Research*, 21:367–384.

Lozano, R., Fullman, N., Abate, D., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., et al. (2018). Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related sustainable development goals for 195 countries and territories: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):2091–2138.

Menken, J. (1987). Proximate determinants of fertility and mortality: A review of recent findings. In *Sociological Forum*, volume 2, pages 697–717. Springer.

Mondal, M. N. I., Hossain, M. K., and Ali, M. K. (2009). Factors influencing infant and child mortality: A case study of rajshahi district, bangladesh. *Journal of Human Ecology*, 26(1):31–39.

Mutunga, C. J. (2004). Environmental determinants of child mortality in urban kenya. In *Discussed in an informal workshop held at the Abdus Salam ICTP: Trieste, Italy*.

Organization, W. H. et al. (2011). Global hiv/aids response: epidemic update and health sector progress towards universal access: progress report 2011.

Organization, W. H. et al. (2015). Global health observatory data repository. 2015. *Substance Use of Mental Health [data file]*.

Root, G. P. (2001). Sanitation, community environments, and childhood diarrhoea in rural zimbabwe. *Journal of health, population and nutrition*, pages 73–82.

Rutstein, S. O. (2005). Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: evidence from the demographic and health surveys. *International Journal of Gynecology & Obstetrics*, 89:S7–S24.

Rutstein, S. O. (2008). Further evidence of the effects of preceding birth intervals on neonatal infant and under-five-years mortality and nutritional status in developing countries: Evidence from the demographic and health surveys.

Skolnik, R. (2008). *Essentials of global health*. Jones & Bartlett Publishers.

Stallings, R. Y. (2004). Child morbidity and treatment patterns.

Titaley, C. R., Dibley, M. J., Agho, K., Roberts, C. L., and Hall, J. (2008). Determinants of neonatal mortality in indonesia. *BMC public health*, 8(1):232.

UNICEF. (2007). *Progress for children: a world fit for children statistical review*. Number 6. Unicef.

UNICEF et al. (2013). Who, world bank, un desa/population division. levels and trends in child mortality—report 2013. estimates developed by the un inter-agency group for child mortality estimation.

Unicef et al. (2015). Unicef data: monitoring the situation of children and women. *The state of the World's children report*.

United Nations, D. o. E. (2006). *The millennium development goals report*. UN.

WHO, U. (2012). Unfpa, the world bank. trends in maternal mortality: 1990 to 2010. who, unicef.

You, D., Bastian, P., Wu, J., and Wardlaw, T. (2013). Estimates developed by the un inter-agency group for child mortality estimation. *Levels and Trends in Child Mortality: Report*.

.1 Appendix

Below are the STATA Do file codes used before being exported to R software

```
keep v001 v002 v024 v025 v113 v190 v208 v714 m15 v116 v149 bord v212 v501 v404 b5

gen alive=.
replace alive=1 if b5==0
replace alive=0 if b5==1
drop b5

egen hhid= concat(v001 v002)
drop v002
rename hhid v002
destring v002,replace

* source of drinking water
recode v113 11/13=1 21/42=2 43/51=3 61/71=1 96/97=4
label define v113 1 "piped" 2 "borehole" 3 "river" 4 "other"
label values v113 v113

* wealth index
recode v190 1=1 2=1 3=2 4=3 5=3
label define v190 1 "poor" 2 "middle" 3 "rich"
label values v190 v190

* marital status
recode v501 0=1 1/2=2 3=3 4/5=4
label define v501 1 "single" 2 "married" 3 "widowed" 4 "divorced"
label values v501 v501

* place of delivery
recode m15 11/12=1 21/31=2 32/33=1 36=2 96=1
label define m15 1 "non-facility" 2 "facility"
```

label values m15 m15

```
* type of toilet facility
recode v116 11/15=1 21/23=2 31/97=3
label define v116 1 "flush toilet" 2 "pit latrine" 3"no facility"
label values v116 v116
```

```
* mother's education
recode v149 0=0 1/2=1 3/4=2 5=3
label define v149 0 "no education" 1 "primary" 2 "secondary" 3 "tertiary"
label values v149 v149
```

```
* birth order
recode bord 1/2=1 3/4=2 5/15=3
label define bord 1 "1-2" 2 "3-4" 3 "5+"
label values bord bord
```

```
* mother's age at birth
recode v212 6/15=1 16/25=2 26/35=3 36/45=4
label define v212 1 "6-15" 2 "16-25" 3 "26-35" 4 "36-45"
label values v212 v212
```

The following are commands used in R for computation.

Empty model 2 with community cluster effect only

$$mod1 < -lme(alive \sim 1, random = \ 1|v001, data = datacomplete20, method = "ML")$$

$summary(mod1)$

Empty model 2 with community cluster effect only

$$mod2 < -lme(alive \sim 1, random = \ 1|v002, data = datacomplete20, method = "ML")$$

$summary(mod2)$

Model 3 with inclusion of predictors with household effect only
$mod3 < -lme(alive \sim as.factor(v024) + as.factor(v025) + as.factor(v113) + as.factor(v116) + as.factor(v149) + as.factor(v190) + as.factor(v501) + as.factor(v212) + as.factor(bord) + as.factor(v714) + v208 + as.factor(v404) + as.factor(m15), random = \sim 1|v001, data = datacomplete20, method = "ML", na.action = na.exclude)$

$summary(mod3)$

Model 4 with inclusion of predictors with community effect only
$mod4 < -lme(alive \sim as.factor(v024) + as.factor(v025) + as.factor(v113) + as.factor(v116) + as.factor(v149) + as.factor(v190) + as.factor(v501) + as.factor(v212) + as.factor(bord) + as.factor(v714) + v208 + as.factor(v404) + as.factor(m15), random = \sim 1|v002, data = datacomplete20, method = "ML", na.action = na.exclude)$

$summary(mod4)$

Model 5 running predictors without cluster effects
$mod5 < -glm(alive \sim as.factor(v024) + as.factor(v025) + as.factor(v113) + as.factor(v116) + as.factor(v149) + as.factor(v190) + as.factor(v501) + as.factor(v212) + as.factor(bord) + as.factor(v714) + v208 + as.factor(v404) + as.factor(m15), data = datacomplete20)$

$summary(mod5)$

MULTILEVEL MODELLING OF FACTORS AFFECTING CHILD MORTALITY IN KENYA

ORIGINALITY REPORT

14%	12%	5%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

4%

★ czep.net

Internet Source

Exclude quotes	On	Exclude matches	< 15 words
Exclude bibliography	On		