



University of Nairobi

**UNIVERSITY OF NAIROBI
COLLEGE OF BIOLOGICAL AND PHYSICAL
SCIENCES
SCHOOL OF MATHEMATICS**

**RISK FACTORS IN CLINICAL MALARIA
MORTALITY AMONG CHILDREN IN KENYA**

BY

**HARMON MUSYOKA MULE
I56/69075/2013**

JULY 2018

**A RESEARCH PROJECT SUBMITTED IN PARTIAL
FULFILMENT OF THE REQUIREMENT FOR DEGREE
OF MASTER OF SCIENCE IN BIOMETRY**

©2018

DECLARATION

This research project is my original work and has not been submitted for examination in any other university or institution of higher learning.

Signature:..... Date:.....

Harmon Musyoka Mule

I56/69075/2013

This thesis has been submitted for examination with my approval as the University Supervisor.

Signature:..... Date:.....

Idah Orowe, Ph.D

School of Mathematics

University of Nairobi

DEDICATION

This research project is dedicated to my loving parents Joshua Mule Mbondo and Rosemary Mutio Mule for their continuous support in my life.

I also dedicate this work to my sisters Edith and Edna, my brother Simon, my nieces Shirley and Hope and my nephew Adriane for the support they accorded to me during my study.

Lastly, I wish to pay tribute to the lecturers who have taught me.

ACKNOWLEDGEMENT

Thanks and appreciation to my supervisor Dr. Idah Orowe for the timely assistance and scholarly guidance accorded to me.

I recognize and appreciate the Ministry of Health in Kenya for provision of the data used in this research project.

Above all, special thanks to God Almighty for giving me good health, wisdom and understanding in doing this research work.

ABSTRACT

Clinical malaria deaths in Kenya is highest among children and therefore there is need to research on the various covariates influencing malaria mortality among children in order to reduce this complex problem that malaria poses in Kenya and sub saharan Africa at large. The main objective of this study was to determine the effect of clinical malaria fatality predictors among children in Kenya. Secondary data was obtained from government hospitals in Kwale, Homabay, Kericho and Busia districts where we have the highest percentage of regional distribution of malaria in children. The data obtained was from January 1999 to December 2012. The findings from this study indicated that there is a linear relationship between malaria mortality in children and predictors malaria admission and rainfall. It was also found that clinical malaria and non-malaria mortalities were significantly different from 1999 to 2012.

Contents

1	INTRODUCTION	1
1.1	Background	1
1.2	Statement of the problem	3
1.3	Research objective	3
1.3.1	General objective	3
1.3.2	Specific objectives	3
1.4	Research methodology	4
1.4.1	Scope and study area	4
1.5	Justification of the study	5
2	LITERATURE REVIEW	6
2.1	Introduction	6
2.2	Review of previous studies	6
3	METHODOLOGY	12
3.1	Introduction	12
3.2	Data collection and organization	12
3.2.1	Data collection	12
3.2.2	Data organization	12
3.3	Data analysis methods	13
3.3.1	Logistic regression	13
3.3.2	The logistic function	14
3.3.3	Logistic regression with several explanatory variables	16
3.3.4	Model fit assessment	17
3.3.5	The wald statistic	17
3.3.6	Likelihood ratio test	18
3.3.7	Hosmer-Lemshow Test	18
3.3.8	Comparison between groups of survival data	18
3.3.9	Test of Hypothesis	19

4	ANALYSIS, RESULTS AND DISCUSSION	21
4.1	Introduction	21
4.2	Analysis of data and results	21
4.2.1	Descriptive analysis of data	21
4.2.2	Inferential Analysis of Data	23
4.3	Discussion	27
5	SUMMARY, CONCLUSION AND RECOMMENDATIONS	29
5.1	Introduction	29
5.2	Findings	29
5.3	Conclusions	29
5.4	Recommendations	30

List of Tables

3.1	The log rank test	19
4.1	Percentage of Malaria admissions and deaths per district . . .	22
4.2	Percentage of Non Malaria admissions and deaths per district	23
4.3	Summary of the model fit	23
4.4	Hosmer and Lemeshow goodness of fit test	24
4.5	Interpretation of the odds ratio	25
4.6	Log Rank test on clinical malaria mortality data	26
4.7	Log Rank test on non-malaria mortality data	27

List of Figures

1.1	Regional distribution of Malaria in children	2
1.2	Study area	4

Chapter 1

INTRODUCTION

1.1 Background

In Kenya, top of the highest causes of morbidity and mortality is Malaria, particularly in children. Malaria is a disease caused by parasites of the genus *Plasmodium*, with *falciparum* being the highest known cause of Malaria. In order to control Malaria within the country, several interventions have been integrated. Among them include providing quick and easily accessible treatment or malaria case management, control of the malaria vector by use of long lasting mosquito nets, prevention and treating malaria during pregnancy and being prepared to respond to the epidemic.

Malaria diagnosis is based on clinical suspicion and detection of parasites in the blood, also referred to as parasitological or confirmatory diagnosis. The current recommendation given by the Ministry of Health in the country is to have diagnosis of malaria which can be confirmed in all groups of patient ages in all epidemiological settings. There are four malaria epidemiological zones within the country, with diversity in risk determined largely by altitude, rainfall patterns and temperature

Globally, the malaria burden has substantially declined as a result of scaling up of effective interventions (Okebe *et al.* 2014). Malaria interventions such as long-lasting mosquito nets, spraying of insecticides indoors, rapid diagnostic tests (RDTs) and artemisinin-based combination therapy (ACT) have played a major role in this decline. Regardless of this substantial decline, malaria continues to be the highest cause of outpatient consultations in Kenya.

Malaria has an uneven distribution with reported determinants of transmission, which include; residential area, overcrowded population, level of education, awareness of malaria prevention, referral status, distance to the nearest health center and socioeconomic factors such as age, sex, parental occupation and income. Globally, demographic factors such as the area of residence contribute largely on disproportionate distribution of malaria cases. Rural areas report a large number of cases as compared to urban settlements (Okebe *et al.*, 2014), attributable to closer proximity to health centers, better housing and drainage systems which reduce mosquito breeding sites.

As reported by Lindblade *et al.* (2004), universal malaria decline in terms of morbidity and mortality has resulted from scale-up of malaria control interventions. They further report that although this decline is an achievement, these current interventions alone will not lead to malaria elimination in most malaria endemic areas, additional strategies must be analyzed and considered.

In Kenya, a multi-prolonged malaria control strategy was introduced, through which Kenya has increased the use of insecticide treated mosquito nets (Noor *et al.*, 2007). Irregardless of the significant steps in the prevention of malaria, the country's search for prompt and effective malaria treatment continues to be a challenge.

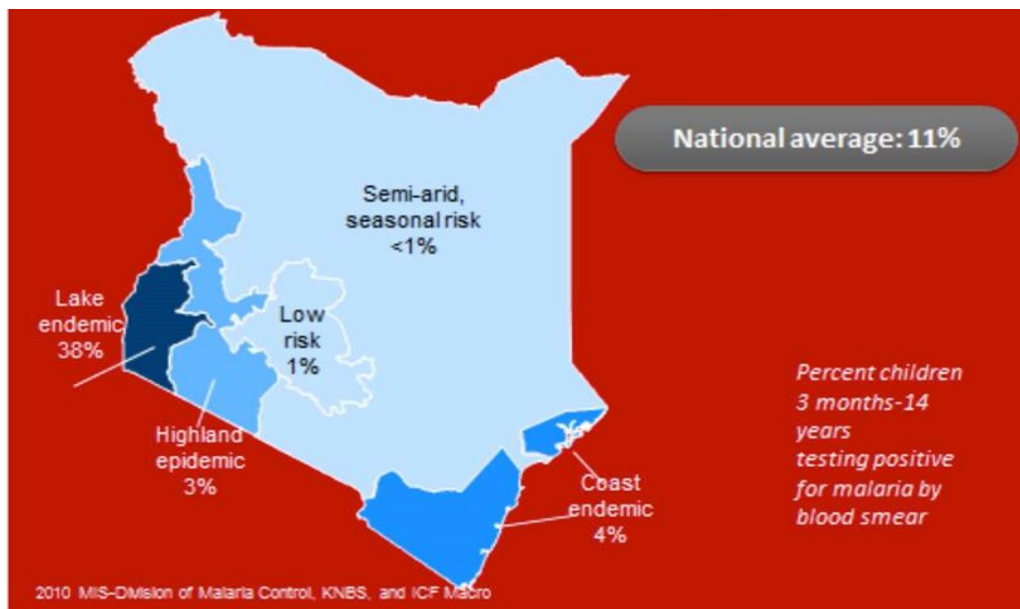


Figure 1.1: Regional distribution of Malaria in children

The lake basin region and surrounding areas experience the highest rate of Malaria infections in children as well as the coastal region of Kenya. The semi-arid areas in Kenya experience the lowest rate of Malaria in children in Kenya. Image adapted and modified from malaria matters.org.

1.2 Statement of the problem

Malaria, a public health concern, causes over one million mortalities globally, and remains the most common type of outpatient consultations in sub saharan Africa despite its global decline over the past few years. Moreover, malaria interventions have led to a change in the risk of infection and disease in as much as they have resulted to decline in disease burden. Furthermore, malaria infections among children in Kenya continues to cause a complex problem in the country, demanding each area in public policy to be mobilized in order to facilitate informative decision making which will see resultant malaria free communities in the country. Therefore there is a need to research on the various covariates influencing malaria mortality among children in Kenya to reduce this complex problem possessed by malaria. It is in this note that this study is aimed at analyzing the effect of clinical malaria fatalities among children in Kenya. Clinical malaria deaths in Kenya is highest among children and therefore there is need to research on the various covariates influencing malaria mortality among infants, to reduce this complex problem malaria poses. It is in this note that this study wishes to analyze the effect of malaria fatality predictors among children in Kenya.

1.3 Research objective

1.3.1 General objective

The main objective of this study is to determine the effect of clinical malaria fatality predictors among children in Kenya.

1.3.2 Specific objectives

1. To examine the significant predictors of clinical malaria fatality among children in Kenya.
2. To apply logistic regression modelling to analyze predictors of clinical malaria fatality among children in Kenya.

3. To apply survival analysis to determine case fatality rates on time of death between malaria admissions and non-malaria admissions.

1.4 Research methodology

1.4.1 Scope and study area

A study done by Chuma *et al.* (2008), in Kenya reported that prompt access to effective malaria treatment remains a challenge for malaria control programs. According to this study done in 2008, evidence shows that prompt access to effective treatment is only available for a minority of the population, with the poorest households being most affected by lack of access. The study also found out that much of malaria seeking occurs in the informal sector. It is in this regard that the study area was chosen away from the major cities in the country.



Figure 1.2: Study area

The highlighted areas represent the study area included in this study. Image adapted and modified from http://d-maps.com/carte.php?num_car=239&lang=en.

The study areas chosen were Kwale, Homabay, Kericho and Busia where clinical malaria is highly reported. From figure 1.2 above, Homabay is an area in the lake basin region while Busia and Kericho are areas surrounding the lake basin region. Kwale is an area in the coastal region of Kenya. In addition, the four areas presented relatively diverse weather and climate conditions which would give this study a broader view of the predictors of clinical malaria mortality in the country.

1.5 Justification of the study

This study is significant to policy makers, stakeholders and the public in general. The findings of this study will add on to the scientific measures and perspective for evaluating malaria mortality especially to children. To the Ministry of Health in Kenya, this study will aid in identifying the risk factors affecting clinical malaria among children, especially in the poorest homesteads. It will also shed some light to whether previous measures set by the government were significant in reducing clinical malaria mortality among children in the country.

To the stakeholders such as investors and employees in the health sector, this study will provide information which will allow them to come up with suggestions on how clinical malaria can be detained. The study will also act as a continuous assessment to the measures being implemented to curb clinical malaria mortality among children in the country.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

A number of researches have been carried out on the application of logistic regression in modelling malaria mortality and other related issues in individuals. A review of literature on these previous researches will be discussed in this chapter.

2.2 Review of previous studies

A study on risk of behaviour for malaria in the resettlement area of Machodinho in the Amazon forests of Brazil as examined by Castilla *et al.* (1993), suggested that economic status and knowledge of the importance and behavior of the mosquito in transmitting malaria are significant factors in determining prevalence risk irrespective of whether preventive precautions, for example, dichlorodiphenyl trichloroethane (DDT) spraying of houses and cleaning of vector breeding sites are to be undertaken in the endemic areas. However, the researchers found out that a higher economic status combined with better knowledge of the vector and DDT spraying of houses decreased the risk of infection. They suggested that a more positive implication is that control programs must work harder and more intensively on behalf of poorer people especially migrants in order to diminish the disease burden for them.

In India, Sharma *et al.* (2001) studied the behaviour of people towards malaria as well as the social and economic factors. This study brought to light the fact that poor social and economic conditions as well as cultural factors have a key role in ensuring malaria is transmitted at a high rate. Of the people-behaviour examined, habits of sleeping, activities people during

the evening and night, the degree of knowledge about malaria treatment are of high importance while determining malaria transmission.

Mensah *et al.* (2004) did a research in Benin to establish malaria transmission in economic factors and how these factors play an important role in aiding transmission of malaria. Their study concluded that majority in the population in Benin had limited knowledge of contributory economic factors to malaria. Ignorance in these factors therefore led to more people acquiring malaria in regions within Benin where Malaria infections were highest.

In Malawi, using spatial modeling of hospital register data, Kazembe *et al.* (2006) examined the patterns of malaria-related hospital admissions and mortality among children in Malawi. With the aim of describing the spatial distribution of children mortality in hospitalization and in-hospitals. This was done through the application of spatial regression models with the aim of aimed spatial variation quantification and risk factors associated with malaria hospitalization and in-hospital mortality. The study used pediatric ward register data from Zomba district in Malawi between years 2002 and 2003 for the case study in which two spatial models were developed, a Poisson model and a logistic model. While adjusting for individual covariates to analyze case fatality rates, to each individual level data, the logistic model was used. On the other hand, the Poisson model was applied to analyze hospitalization and minimum mortality rates with sex and age as covariates. From this study, it was concluded that the malaria hospitalization and in-hospital mortality rates decreased with age. In addition, the rate of death was highest on day one then subsequent days had low rate as length of hospital stay increased in days.

De La Cruz *et al.* (2006) made use of logistic regression modelling and other tests to identify factors associated with bed net use among children aged not more than five years in order to compare the characteristics of caregivers whose children use mosquito nets to those whose children do not use mosquito nets. The study, carried out in Ghana, identified factors and characteristics of mothers that affect use of mosquito nets among their children aged less than five years. Data was obtained from the baseline component of an evaluation of Freedom from Hungers malaria curriculum. To select the clients, a quasi-experimental design was used to select 516 clients and 535 non clients. Region of residence, greater food security, caregivers beliefs about symptoms, causation and groups most vulnerable to malaria were factors considered most closely associated with use of mosquito nets. The study concluded that increased knowhow about malaria may not always result into

improved mosquito net use.

Ye *et al.* (2007) assessed the self-diagnosed malaria risk in urban informal settlements of the capital city of Kenya by use of a morbidity survey that was self-reported. Nairobi is a low risk area for malaria. However, with the aim of exploring perceived malaria risk and some associated factors, the survey was carried out in two informal settlements within Nairobi from May to August 2004 on 7288 participants. During the survey, done in Viwandani and Korogocho, individuals were requested to report illnesses they experienced within the last two weeks. The study used logistic regression to estimate the odds of perceived malaria. The survey revealed that participants in the study reported 165 illnesses, with malaria being the leading illness at 28.1%. Perceived malaria risk was found to be significantly higher in Viwandani as compared to Korogocho with an odds ratio of 1.61 and confidence interval of 95% CI (1.10 2.26). Participants in age group of 25-39 years had significantly higher odds of perceived malaria compared to those under five years with an odds ratio of 2.07 and a 95% confidence interval of 1.43-2.98. With respect to ethnicity, the Kikuyu had reduced odds of perceived malaria compared to other ethnic groups. The results from this survey necessitated for a further assessment of malaria epidemiology in Nairobi.

In Zambia, Riedel *et al.* (2010) made use of logistic regression models to develop geographical patterns and predictor of malaria risk. In 2006, the Zambian Ministry of Health in conjunction with government partners in Zambia carried out a nation-wide malaria survey, Zambia Malaria Indicator Survey (ZMIS). The aim of this study was to estimate the coverage of interventions and malaria related burden in children less than five years. One of the objectives of this study was to estimate an empirical high-resolution parasitological risk map in the country through analysis of the data obtained. Secondly, ZMIS wanted to assess the relation between malaria interventions and parasitaemia risk after adjusting for environmental and socio-economic confounders. In order to assess the relationship between the parasitaemia risk outcome and the environmental lag time variables, the study fitted bivariate logistic regression models. The study established that for each climate factor, the lag time variable gave a model with the smallest Akaike's Information Criterion (AIC). In a 15% level of significance determined by likelihood ratio tests, all covariates which were significant in the bivariate analysis were included in a multiple geo-statistical logistic regression analysis. The study further fitted several geo-statistical multiple logistic regression models to assess and capture potential non-linearity in the malaria-environment relation. These models included covariates in continuous scales, quartile categorical

scales and fitted by penalized and basis spline (P- and B- splines) curves. The model with the best predictive ability was then chosen through model validation procedure. The covariates and random effects were modelled on the logit scale of the parasitaemia risk parameters, with their geo-statistical models having at least as many parameters as the number of locations. Non-spatial models having smaller number of parameters were fitted as a result of exploratory analysis done which suggested weak spatial correlation. Model validation indicated that linear environmental predictors were able to fit the data as well as or even better than more complex non- linear terms and that the data do not support spatial independence. One of the conclusions of the study was that the averaged was 20.0% in children less than five years.

In order to study the probability of infection in severely ill patients, Daliana and colleagues studied Infection Probability Score with the objective of developing a simple score to assist determine the presence or absence of infection in severely ill patients by use of readily available variables. The study used measurements from temperature of the body, heart rate, rate of respiration, count of white blood cells and C-reactive protein concentrations with infection presence defined using the CDC definition. The SOFA (i.e. Sequential Organ Failure Assessment) score was calculated throughout the ICU stay. Infection was found in 26% and 29% of the patients in the developmental and validation sets respectively. In order to select predictors which were significant to infection, univariate logistic regression was used. This was done by transforming each predictor which was continuous into a variable which was categorical by use of weighted least square regression. The variables were coded 1 or 0 to allow for scoring of the different predictors relatively, a multiple logistic regression model was fitted to predict infection. The fitted model brought to light that the significant variables for infection were rate of heart beat and C-reactive protein while rate of respiration was found to be the least significant variable.

In Bangladesh, Reid *et al.* (2010) developed mapping malaria risk using Bayesian geo-statistical models. In order to understand better the factors affecting malaria transmission while accurately determining the limits of malaria transmission, researchers have made advancements in geo-statistical methods and geographical information systems (GIS). A prevalence survey of malaria was carried out in Bangladesh in 2007 with N=9,750 individuals and N=354 communities. The data obtained was then standardized to a range of age of between two to less than ten years. Bayesian geo-statistical logistic regression models with environmental covariates predicted prevalence of *Plasmodium falciparum* for children aged between 2 to 10 years i.e PfPR₂₋₁₀ in

the areas of Bangladesh which were endemic. An estimation of the number of individuals living in different endemic classes was done by combining the predictions with gridded population data. In addition, environmental variables across the endemic areas included vegetation coverage, elevation and minimum temperature. The study concluded that the final Bayesian geostatistical model had the good ability to predict. The study recommended that contemporary GIS and geostatistics that were model-based could be made to use in interpolation of malaria risk in Bangladesh.

It can be difficult to establish parasitaemia in malaria diagnosis for areas where malaria is most prevalent. Therefore, Essuman *et al.* (2010) did a study on severe malaria retinopathy in children in Ghana between variations in cerebral and non-cerebral malaria. Their research was intended, first of all, to determine use of diagnostic retinopathy on ophthalmoscopy in severe malaria syndromes, that is, respiratory distress malaria and severe anaemia malaria, among children in Ghana. Further, it was aimed at determining the association between retinopathy and the convulsion occurrence in patients with cerebral malaria. A cross-sectional research of consecutive admitted severe malaria patients were tested for signs of retinal. The study found that whitening of the retinal, which was a sign of ischaemia of the retinal, was statistically significant more common in cerebral malaria than in non cerebral malaria syndromes.

In 2013, Bonface did a retrospective review of patients medical records to abstract data on diagnosis for those who died in year 2013. The study objective was to find out the mortality rate and clinical profiles of children aged 5-17 years who died in six Kenyan hospitals in the year 2013. Through use of descriptive statistics, data from the six Kenyan hospitals which explored differences in mortality rates between age groups and gender was analysed. The review revealed that out of 4,520 patients records retrieved, the in-hospital mortality rate was 3.5% (95%CI 3.0-4.1) with variations in deaths between the ages and gender. Among the deaths, 60% suffered from communicable diseases, maternal and nutritional causes; 41.3% suffered from non-communicable diseases. A further 11.9% succumbed to traumatic injuries. The predominant clinical diagnoses among patients who died were HIV/AIDS, respiratory tract infections and malaria. It was therefore concluded that infectious causes had the highest proportion of diagnoses among children aged 5-17 years who died in Kenya.

Abdur *et al.* (2017) studied malaria prevalence to find out the associated factors in social and economic status among Kenyan children from

ages of between 6 months to 14 years. The data used in this study was from KMIS 2015 survey. Multivariate logistic regression analysis was used to analyse demographic associations, social and economic status, community and behavioral factors with malaria prevalence among children in Kenya. The study found that malaria prevalence had an upward trend with respect to age, with the highest prevalence among children aged 11 to 14 years. Malaria Prevalence was also higher among rural children (10.16%) compared to urban children (2.93%), as well as poor children (11.05%) compared to rich children (3.23%). The likelihood of having malaria was higher among children aged 10-14 years ($AOR = 4.47, 95\%CI = 3.33; 6.02; P < 0.001$) compared with children aged under 5 years. The presence of anemia ($AOR = 3.52, 95\%CI = 2.78; 4.45; P < 0.001$), rural residence ($AOR = 1.71, 95\%CI = 1.31; 2.22; P < 0.001$), lack of a hanging mosquito net ($AOR = 2.38, 95\%CI = 1.78; 3.19; P < 0.001$), primary education level of the household head ($AOR = 1.15, 95\%CI = 1.08; 2.25; P < 0.05$), and other factors, such as the household having electricity and access to media such as television or radio, were also associated with the likelihood of infection. This study demonstrated the need to focus on awareness programs to prevent malaria and to use existing knowledge in practice to control the malaria burden in Kenya. Furthermore, this study suggests that improving the information available through the mass media and introducing behavior change communication and intervention program specifically for those of poor socioeconomic status will help to reduce malaria cases.

Chapter 3

METHODOLOGY

3.1 Introduction

This chapter discusses data collection method, how the data is organized and presented for data analysis. This chapter will also elucidate the key statistical techniques that will be used for analysis of the data.

3.2 Data collection and organization

3.2.1 Data collection

The data used in this study was obtained as secondary data from government hospitals in four regions in Kenya namely Kwale, Homabay, Kericho and Busia from 1st January 1999 to 31st December 2012. Each admission case for children below the age of 15 years was clinically assessed and confirmations of malaria admission or non malaria admissions were done. For each fatality from any patient admitted who was below 15 years of age, the cause of death was registered as either due to malaria or non-malaria. The register also included rainfall received within the month and the maximum and minimum temperature for the month was also recorded.

3.2.2 Data organization

From the secondary data obtained, the following variables were derived: outcome (dead due to clinical malaria = 1 and survival after admission due to clinical malaria = 0); year (1999 to 2012); month of the year (January =1 up to December =12); number of clinical malaria admissions within the month; number of non-malaria admissions within the month; non-malaria mortality

within the month; rainfall received within the month in mm; average temperature within the month.

The data obtained was analyzed using R program.

3.3 Data analysis methods

3.3.1 Logistic regression

In this study, logistic regression modeling was used to estimate the probability that a child will fall into either dead i.e. malaria fatality which is the event of interest, or recovered. This probability would in turn aid in interpretation of the coefficients in the logistic regression model.

The study focused also on the odds ratios. We were interested in the probability of a malaria fatality occurring. Therefore, P was defined as the probability of a Malaria fatality occurring. Hence, the probability of recovering from Malaria automatically becomes $(1 - P)$. Given this information, the odds of a Malaria fatality occurring would now be defined as the ratio of the probability that a Malaria fatality would occur versus the probability that a recovery would occur.

$$Odds = \frac{p}{1-p} \quad (3.1)$$

This study aimed at considering several predictors in the determination of whether a Malaria fatality would occur or not (recovery). Therefore the odds ratio was calculated separately for each predictor. The odds ratio in this study referred to the measure of the odds of Malaria fatality occurring in one group compared to the odds of Malaria fatality occurring in another group.

However, the joint effect of all the predictors used in this study was expressed as follows:

$$Odds = \frac{p}{1-p} = e^{\alpha + \sum_{i=1}^k \beta_i x_i} \quad (3.2)$$

In order to determine the significance of each of the covariates in the above expression (i.e. x_1, x_2, \dots, x_k) to the overall odds ratio, the logarithm

of both sides of the equation was taken.

Assumptions of logistic regression model

- i The response variable is binary in nature
- ii Logistic regression requires large samples to be accurate

3.3.2 The logistic function

The logistic model is based on the logistic function. The logistic function is given as;

$$f(x) = \frac{1}{1+e^{-x}} \quad (3.3)$$

On plotting the values of this function, x varies from $-\infty$ to $+\infty$. This means that the range of $f(x)$ is always between 0 and 1 regardless of the value of x .

The logistic model

The response variable in the logistic regression is binary. It either takes the value of 1, i.e. probability of success (p), or it takes the value 0, the probability of failure ($1 - p$). A variable which takes either these two values is called a Bernoulli variable.

The response variable takes the binary form, while the predictor variables in any logistic regression may take any form. The logistic regression model makes no assumption on the distribution of the predictor variables. The predictor variables need not be normally distributed. The relationship between the predictor variables and response variable does not need to be a linear function, however, the logistic regression function used is the logit transformation of p .

The transformation was done by writing x as the linear sum:

$$x = \alpha + \sum_{i=1}^k \beta_i z_i \quad (3.4)$$

In the above equation, z_i represented the independent variables while α and β_i were the unknown constant parameters, where $i = 1, 2, \dots, k$. The logistic equation would then be written as:

$$f(x) = \frac{1}{1+e^{-(\alpha+\sum_{i=1}^k \beta_i x_i)}} \quad (3.5)$$

Now, let $p(x)$ be the probability statement where x runs from x_1 through to x_k . Therefore the logistic model was expressed as follows:

$$p(x) = \frac{1}{1+e^{-(\alpha+\sum_{i=1}^k \beta_i x_i)}} \quad (3.6)$$

Logit transformation then was used to make the above non-linear model to be linear. Introducing the logit transformation resulted to:

$$\text{logit}[p(x)] = \ln \frac{p(x)}{1-p(x)} \quad (3.7)$$

where $p(x) = \frac{1}{1+e^{-(\alpha+\sum \beta_i x_i)}}$

After this transformation, it was then possible to compute the $\text{logit}p(x)$ for an individual with independent variables given by x .

$\ln \frac{p(x)}{1-p(x)}$ was then represented in terms of $p(x)$.

$$\ln \frac{p(x)}{1-p(x)} = \ln \frac{\frac{1}{1+e^{-(\alpha+\sum \beta_i x_i)}}}{\frac{e^{-(\alpha+\sum \beta_i x_i)}}{1+e^{-(\alpha+\sum \beta_i x_i)}}} \quad (3.8)$$

$$= \ln[e^{\alpha+\sum \beta_i x_i}] \quad (3.9)$$

$$= \alpha + \sum \beta_i x_i \quad (3.10)$$

The logit of $p(x)$ was simplified to a linear sum.

$$\text{logit}p(x) = \alpha + \sum \beta_i x_i \quad (3.11)$$

Now, given that the log of $\frac{p(x)}{1-p(x)}$ gives the logit, this logit describes the odds for a malaria fatality occurring, with the predictor variables specified by x .

Logistic regression

An S-shaped curve can be transformed into an approximate straight line by use of the logistic function. This function can also be used to change the range of a random variable from 0 to 1 to $-\infty$ to $+\infty$.

The logistic function is defined as the natural logarithm (\ln) of the odds of an event. We then represented the logistic function as:

$$\text{logit} = \ln \frac{p}{1-p} \quad (3.12)$$

The underlying distribution is binomial with parameters α and β . These parameters were estimated using the method of maximum likelihood.

In this study, it was our wish to investigate how the fatality of malaria patients can be predicted by several predictor variables. These predictor variables may be either qualitative or quantitative. A stepwise method was used to test for the signi-

cance of inclusion or elimination of the predictor variables at each stage, based on the change in likelihood, or using the p-value and other test statistics.

3.3.3 Logistic regression with several explanatory variables

In this study, we required the logistic regression parameters estimated to be in a way that the coefficients made the observed results more likely. Therefore the use of Maximum Likelihood Estimators (MLEs) would be necessitated. The maximum likelihood estimation would involve finding the value of the parameter that would give rise to the maximum likelihood. A likelihood is the probability that the number of successes in a sample of size n is exactly r .

Let $f(x_1, x_2, \dots, x_n; \theta)$ be the joint probability (or density) function of n random variables x_1, x_2, \dots, x_n . The likelihood function of the sample is given by $L(\theta; x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n; \theta)$. Maximizing the likelihood function $L(\theta)$ is equivalent to maximizing the natural logarithm of $L(\theta)$ i.e. $\ln L(\theta)$ which is computationally easier.

3.3.4 Model fit assessment

The first step in logistic regression analysis is estimating the coefficients of the predictor variables. However, there was need to further test the appropriateness, adequacy and usefulness of the model. This is done by testing the significance of each predictor variable and thereafter the overall goodness of fit of the model.

The following methods are useful in testing the significance of the predictor variables.

3.3.5 The wald statistic

The statistical significance of each coefficient in the model would be tested by using the wald statistic. The wald statistic has a chi-square distribution shown below:

$$z^2 = x^2 = \left(\frac{\beta}{SE}\right)^2 \quad (3.13)$$

where $z = \frac{\beta}{SE}$

The MLE of the parameters of interest under the Wald statistic was compared to the proposed value. However, this assumed that the difference between the two would approximately be normally distributed. The square of this difference is what was compared to the chi-square distribution with 1 degree of freedom.

It is important to note that when there is only one regression coefficients of interest under interval estimation, a large sample confidence interval for the parameter is obtained. This is done by computation of the estimate of the parameter plus or minus a percentage point of the normal distribution multiplied by the estimated standard error.

$$100(1 - \alpha)\% \text{ CI for } \beta_i \quad (3.14)$$

CI for Odds Ratio was given by:

$$e^{(CI \text{ for } \beta_i)} \quad (3.15)$$

Therefore, at 95% confidence interval with exposure variable X_i , lying within 0 and 1, the adjusted odds ratio would be given by:

$$e^{(\beta_i + 1.96 * s\beta_i)} \quad (3.16)$$

3.3.6 Likelihood ratio test

The likelihood ratio test was used to compare the goodness of fit between the null model (L_0) and the alternate model (L_1).

The test statistic for the likelihood ratio test was calculated as follows:

$$-2 * \ln(\text{likelihoodratio}) \quad (3.17)$$

$$-2 * \ln\left(\frac{\text{likelihood for null model}}{\text{likelihood for alternative model}}\right) \quad (3.18)$$

$$-2 * \ln\left(\frac{L_0}{L_1}\right) \quad (3.19)$$

$$-2 * (\ln L_0 - \ln L_1) \quad (3.20)$$

This statistic would then be compared with a chi-square distribution with 1 degree of freedom.

3.3.7 Hosmer-Lemshow Test

The HosmerLemeshow test was used as a test to assess the goodness of fit of the model. This test was used since it would allow for any number of predictor variables, which would be either categorical or continuous. In this study, this test, which is similar to a X^2 goodness of fit test, had the advantage of partitioning the observations into 10 ordered groups of subjects groups which were of equal size approximately. The observations were then grouped into deciles based on the predicted probabilities. The test statistic was calculated, based on the data used for this study, using the expected and observed counts, and had an approximate X^2 distribution with 8 (=10 - 2) degrees of freedom.

3.3.8 Comparison between groups of survival data

To perform a comparison between groups of survival data, the Log Rank test would be performed. The hypothesis below is tested using this test:

H_0 : There is no difference between the groups of survival data

H_1 : There is a difference between the groups of survival data

Log Rank Test

The log rank test considered separately each death time in two groups: say group I and group II, where the death times were independent of one another.

Group	#deaths	#at risk(just before t_i)	#surviving (beyond t_i)
I	d_{1j}	r_{1j}	$r_{1j} - d_{1j}$
II	d_{2j}	r_{2j}	$r_{2j} - d_{2j}$
Total	d_j	r_j	$r_j - d_j$

Table 3.1: The log rank test

Assuming the totals were fixed, then d_{1j} would be a random variable (which can take the values 0 to $\min(r_{1j}$ and d_j) and would have a hyper geometric distribution. Hence the expected value of d_{1j} was the expected number of people who die at time t_j in group I.

$$E(d_{1j}) = e_{1j} = \frac{r_{1j}d_j}{r_j} \quad (3.21)$$

and therefore

$$Var(d_{1j}) = V_{1j} = \frac{r_{1j}r_{2j}d_j(r_j - d_j)}{r_j^2(r_j - 1)} \quad (3.22)$$

Now, given that;

$$U_L = \sum_{j=1}^r d_{1j} - e_{1j} \quad (3.23)$$

where

$$E(U_L) = 0$$

then

$$Var(U_L) = V_L = \sum_{j=1}^r \frac{r_{1j}r_{2j}d_j(r_j - d_j)}{r_j^2(r_j - 1)} \quad (3.24)$$

3.3.9 Test of Hypothesis

In order to learn more about the behaviour in populations that are often too large or inaccessible, inferential statistics was used. The samples used were related to populations. Hypothesis testing is the method in which selected

samples are used to learn more about characteristics in a given population, with the goal being to determine whether a population parameter is likely to be true. There are four main steps of hypothesis testing which have been used in this study, which include stating the hypothesis, setting the criteria for a decision, computing the test statistic and making a decision.

Null and alternative hypothesis

The null hypothesis H_0 stated the assumption to be tested. The alternative hypothesis H_1 was the hypothesis that we believed to be true. The p-value, which was widely used in hypothesis testing, was the probability of obtaining a sample outcome, given that the value stated in the null hypothesis was true. The p-value was compared to the level of significance chosen.

Chapter 4

ANALYSIS, RESULTS AND DISCUSSION

4.1 Introduction

This chapter covers data presentation, data analysis and interpretation. It presents the analysis of the research findings. Data analysis is aimed at answering the research questions of the study which sought to:

- To determine the effect of clinical malaria fatality predictors among children in Kenya.
- To examine the significant predictors of clinical malaria fatality among children in Kenya.
- To apply logistic regression modelling to analyze predictors of clinical malaria fatality among children in Kenya.
- To apply survival analysis to determine case fatality rates on time of death between malaria admissions and non-malaria admissions.

4.2 Analysis of data and results

4.2.1 Descriptive analysis of data

In Busia, it was observed that out of 26,917 malaria admissions between the year 1999 and the year 2012, the highest number of malaria admissions was in the year 2008 with 3,039 admissions (11.3%), with the same year recording the highest number of malaria mortality of 153 deaths (14.9%). Within the same period in Busia, it was also observed that the highest number of

malaria admissions was in the month of May with 2,940 admissions (10.9%) the same month recording the highest number of malaria mortality.

In Homabay, it was observed that out of 19,020 malaria admissions between the year 1999 and the year 2012, the highest number of malaria admissions was in the year 2005 with 1,999 admissions (10.5%), with the previous year recording the highest number of malaria mortality of 182 (10.2%). Within the same period in Homabay, it was also observed that the highest number of malaria admissions was in the month of January with 2,006 (10.5%) admissions and July recording the highest number of malaria mortality.

In Kericho, it was observed that out of 18,587 malaria admissions between the year 2000 and the year 2012, the highest number of malaria admission was in the year 2003 with 2,608 admissions (14%), with the previous year recording the highest number of malaria mortality of 111 deaths (14%). Within the same period in Kericho, it was also observed that the highest number of malaria admissions was in the month of July with 2,389 (12.9%) admissions and the same month recording the highest number of malaria mortality.

In Kwale, it was observed that out of 8,387 malaria admissions between the year 1999 and the year 2012, the highest number of malaria admission was in the year 1999 with 1,053 (12.6%) admissions, with the year 2007 recording the highest number of malaria mortality of 81 (1.8%) deaths. Within the same period in Kwale, it was also observed that the highest number of malaria admissions was in the month of July with 799 (9.5%) admissions with the same month recording the highest number of malaria mortality.

District	Malaria admissions	Malaria deaths
Busia	36.9%	24.7%
Homabay	26.1%	43.2%
Kericho	25.5%	17.8%
Kwale	11.5%	14.2%

Table 4.1: Percentage of Malaria admissions and deaths per district

From table 4.1 above, it was noted that the district with the highest percentage of malaria admissions did not have the highest number of malaria deaths.

District	Non-malaria admissions	Non-malaria deaths
Busia	28.8%	24.7%
Homabay	19.9%	24.6%
Kericho	30.5%	32.7%
Kwale	20.8%	18.0%

Table 4.2: Percentage of Non Malaria admissions and deaths per district

4.2.2 Inferential Analysis of Data

From table 4.3, the logistic model was obtained as follows;

$$(P(y = 1)) = 1.768460 + 0.0075821 \text{Malaria Admissions} - 0.005074 \text{NonMalaria Admissions} - 0.002850 \text{Rainfall} + 0.033870 \text{Average Temperature}$$

The hypothesis below was tested:

H_0 : The predictors used are not significant to the model

H_1 : The predictors used are significant to the model

Coefficients	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.768460	0.358561	4.932	8.14e-07
MalariaAdmissions	0.007582	0.002382	3.183	0.00146
NonMalariaAdmissions	-0.005074	0.003532	-1.437	0.15085
Rainfall	-0.002850	0.001324	-2.154	0.03128
AverageTemperature	0.033870	0.018405	1.840	0.06573

Table 4.3: Summary of the model fit

It can be noted that predictors Malaria admissions and rainfall amount were found to have significance values of 0.00146 and 0.03128 respectively each which are less than $\alpha=0.05$. Therefore the null hypothesis (H_0) is rejected. The conclusion made is that there is enough evidence to show that these predictors are each not equal to zero at 95% confidence level. Therefore these predictors are each important to the final model and hence should be included. Further, these predictors are relevant in predicting malaria mortality in Kenya from 1999 to 2012.

Models not in the model explained

The predictors Non malaria admissions and Average temperature were dropped from the model since the p-values 0.15085 and 0.06573 were each greater than

$\alpha=0.05$. We therefore fail to reject the null hypothesis and conclude that there exists sufficient evidence to indicate that the predictors Non malaria admissions and average temperature are each equal to zero. This means that the predictors Non malaria admissions and average temperature are not relevant in predicting malaria mortality in Kenya from 1999 to 2012.

Model Fit Assessment by Hosmer and Lemeshow Goodness of Fit Test

H_0 : The hypothesized data fits the model

H_1 : The hypothesized data does not fit the model

The model fitted was tested for goodness of fit using the Hosmer and Lemeshow test. The table below shows the findings of this test:

Chi-square	df	P-value
8.0644	8	0.4272

Table 4.4: Hosmer and Lemeshow goodness of fit test

Since the p-value, 0.4272 is greater than the significance level, $\alpha=0.05$, we fail to reject the null hypothesis (H_0) and conclude that there is enough evidence to show that the hypothesized model fits the data set used in predicting malaria mortality in Kenya from 1999 to 2012. This in essence confirms that the number of clinical malaria mortality are not significantly different from those predicted by the model and the overall model fit is good.

Given that the rainfall amount was found to be significant in predicting clinical malaria mortality in Kenya while average temperature was not found to be significant in predicting clinical malaria mortality in Kenya, the study sought to assess the dependence of rainfall amount to average temperature using Pearson Chi-Square test on the dependence of rainfall amount to average temperature. The null and alternative hypothesis for assessing the dependence of rainfall to average temperature;

H_0 : Rainfall amount is independent of the average temperature

H_1 : Rainfall amount is dependent on the average temperature

The p-value was found to be 0.05495 which is greater than the significance level $\alpha=0.05$. We therefore fail to reject the null hypothesis (H_0) and conclude that there is sufficient evidence to show that rainfall amount is independent of the average temperature.

The study sought to find out whether the significant predictors in the model, i.e. malaria admissions and rainfall were dependent on each other. The null and alternative hypothesis for assessing the dependence of malaria admissions to rainfall amount was given by

H_0 : Malaria admission is independent of the rainfall amount

H_1 : Malaria admission is dependent on the rainfall amount

The p-value was found to be 0.00009 which is less than the significance level $\alpha=0.05$. We therefore reject the null hypothesis (H_0) and conclude that there is sufficient evidence to show that malaria admission is dependent on rainfall amount.

The study also sought to find out whether non malaria admissions and rainfall were dependent on each other. The null and alternative hypothesis for assessing the dependence of non malaria admissions to rainfall amount was given by

H_0 : Non-Malaria admission is independent of the rainfall amount

H_1 : Non-Malaria admission is dependent on the rainfall amount

The p-value was found to be 0.09092 which is greater than the significance level $\alpha=0.05$. We therefore fail to reject the null hypothesis (H_0) and conclude that there is sufficient evidence to show that non-malaria admission is independent of rainfall amount.

Interpretation of the Odds ratio

From table 4.5 below, the predictor Malaria admissions recorded an odds ratio of 1.0076103. This indicated that for patients below the age of 15 years who had been admitted due to clinical malaria were 0.7% more likely to die of clinical malaria as compared to those who had not been admitted due to clinical malaria, controlling for all other factors in the model. The odds ratio for rainfall was 0.9971537. This indicated that the rainfall received was 0.28% less likely to cause clinical malaria fatality as compared to areas that did not receive rainfall, controlling for all other factors in the model.

Coefficient	Odds Ratio
Malaria Admissions	1.0076103
Non Malaria Admissions	0.9949393
Rainfall	0.9971537
Average Temperature	1.0344503

Table 4.5: Interpretation of the odds ratio

Survival analysis on case mortality rates between clinical malaria and non-malaria deaths

The study sought to find out the case mortality rates between clinical malaria and non-malaria deaths between the year 1999 to 2012. The first hypothesis tested was that of clinical malaria survival data.

H_0 : There is no significant difference among the groups of malaria survival data in the years between 1999 to 2012

H_1 : There is a significant difference between the groups of malaria survival data in the years between 1999 to 2012

The test was carried out using the Log Rank non-parametric method test.

Log Rank test on clinical malaria mortality data

Year	N	Observed	Expected	$(O - E)^2/E$	$(O - E)^2/V$
1999	36	21	17.4	0.735	0.7685
2000	48	37	26.4	4.3043	4.5838
2001	48	45	28.0	10.2540	10.9458
2002	48	41	30.0	3.9934	4.2764
2003	48	47	32.2	6.8046	7.3103
2004	48	47	34.5	4.5132	4.8666
2005	48	44	37.1	1.2749	1.3805
2006	48	42	39.8	0.1181	0.1285
2007	48	46	42.7	0.2526	0.2763
2008	48	45	46.2	0.0292	0.0322
2009	48	47	50.4	0.2277	0.2537
2010	48	46	55.8	1.7128	1.9446
2011	48	37	62.1	10.1557	11.9384
2012	48	32	74.3	24.0985	33.3946

Table 4.6: Log Rank test on clinical malaria mortality data

The p-value was found to be $1.65e^{-12}$ with a Chisquare value of 84.4 on 13 degrees of freedom which is less than the significance level, $\alpha=0.05$. We therefore reject the null hypothesis H_0 and conclude that there is a significant difference between the groups of malaria survival data in the years between 1999 and 2012.

The second hypothesis tested was that of non-malaria survival data.

H_0 : There is no significant difference among the groups of non-malaria survival data in the years between 1999 to 2012

H_1 : There is a significant difference between the groups of non-malaria sur-

vival data in the years between 1999 o 2012

The test was carried out using the Log Rank non-parametric method test.

Log Rank test on non-malaria survival data

Year	N	Observed	Expected	$(O - E)^2/E$	$(O - E)^2/V$
1999	36	23	19.3	0.6949	0.7268
2000	48	39	28.7	3.6612	3.9002
2001	48	45	30.5	6.8351	7.2985
2002	48	42	32.6	2.6906	2.8815
2003	48	48	34.8	4.9988	5.3697
2004	48	48	37.2	3.1087	3.3513
2005	48	47	39.9	1.2734	1.3783
2006	48	48	42.8	0.6443	0.7006
2007	48	47	46.0	0.0229	0.0250
2008	48	48	49.6	0.0544	0.0599
2009	48	48	54.0	0.6657	0.7414
2010	48	48	59.5	2.2058	2.5001
2011	48	47	66.8	5.8545	6.8765
2012	48	48	84.2	15.5854	22.6034

Table 4.7: Log Rank test on non-malaria mortality data

The p-value was found to be $4.89e^{-8}$ with a Chisquare value of 60.2 on 13 degrees of freedom which is less than the significace level, $\alpha=0.05$. We therefore reject the null hypothesis H_0 and conclude that there is a significant difference between the groups of non-malaria survival data in the years between 1999 and 2012.

4.3 Discussion

This study provided evidence of the predictors of clinical malaria mortality among children aged 0-15 years in Kenya. The data collected was from four regions scattered within Kenya, namely Kwale, Homabay, Kericho and Busia. The model indicates that the number of malaria admissions and the amount of rainfall received within the region are significant predictors of clinical malaria mortlaity among children in Kenya.

While it is commonly perceived that both rainfall and temperature would have an impact on the clinical malaria mortality in Kenya, the model indicates that temperature does not a significant predictor of clinical malaria mortality. The study also found out that rainfall amount in a region in Kenya would not be dependent on the average temperature. This is clearly evidenced in the weather patterns of distinct areas within Kenya. For instance, the Mt. Kenya region which receives high rainfall has generally low temperatures while the lake region of the country which also receives high and regular rainfall generally experiences high temperatures.

The study provided evidence that the number of malaria admissions was dependent on the amount of rainfall. This could be brought about by the poor drainage systems within the country. Therefore as the rainfall amount increases, more water clogs are formed. The water clogs are well known to be breeding areas for mosquitos which in turn transmit malaria.

Survival analysis was used in the study to find out whether there were any significant differences within the groups of clinical malaria mortality and non malaria mortality among the 14 years from 1999 to 2012. It was established that each year was significantly different from the other. This could be explained by the ever changing weather patterns in Kenya, where some years experienced severe drought hence minimal rainfall, while other years experienced heavy rainfall downpour.

It is worth noting that the analysis conducted in this study was based on data collected in four regions within Kenya. However, most studies done in Africa, Kenya included, indicate that majority of malaria treatments and deaths occur outside the hospitals and may therefore fail to be recorded by the relevant authorities.

Chapter 5

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

5.2 Findings

Five key findings from this study were arrived at:

- The predictors malaria admissions and amount of rainfall were found to have an influence on clinical malaria mortality among children in Kenya
- The number of malaria admissions was found to be dependent on the amount of rainfall
- The number of non malaria admissions was found to be independent of the amount of rainfall
- The amount of rainfall was found to be independent of the average temperature
- Clinical malaria and non-malaria mortality were significantly different from 1999 to 2012

5.3 Conclusions

The findings of this study indicates that there is a linear relationship between malaria mortality among children in Kenya and the predictors number of

malaria admissions and amount of rainfall from years 1999 and 2012. The overall logistic model obtained was:

$$(P(y = 1)) = 1.768460 + 0.0075821 \text{Malaria Admissions} - 0.005074 \text{NonMalaria Admissions} - 0.002850 \text{Rainfall} + 0.033870 \text{Average Temperature}$$

The study, from the test of significance of coefficients, found that the predictors average temperature and non malaria admissions were not good predictors of clinical malaria mortality. It was also found that the number of clinical malaria admissions were dependent on the amount of rainfall received.

Through survival analysis on the clinical malaria mortality data, the study also found that there was a significant difference in the clinical malaria mortality rates between the years 1999 to 2012.

5.4 Recommendations

- Based on the findings and conclusions arrived from this study, the Ministry of Health in Kenya should develop a strategic plan to ensure early diagnosis of clinical malaria and therefore treat it before patient admission is necessitated. This could reduce the clinical malaria mortality rate.
- Secondly, stakeholders such as World Health Organization should continuously monitor and analyze data of all illnesses from all areas within the country to gain a better understanding of the risk factors for each disease.
- Finally, further research may be done in this area by considering other factors that may be deemed significant in predicting clinical malaria mortality.

REFERENCES

- Castilla R.E. and Sawyer D.O., (1993). Malaria Rate and Fate: A Socioeconomic Study of Malaria in Brazil. *Social Sciences Medical Journal*, pp 37.
- Sharma. S. K., Pradham. P., Pradhi. D. M., (2001). Socioeconomic Factors Associated with Malaria in a Tribal Area of Orissa, India. *Indian Journal of Public Health*, pp 93-98.
- Mensah, O. and Kumaranyake L. (2004). Malaria Incidence in Rural Benin: Does Economic Matter in Benin Health Policy. *Health Policy Journal*. Volume 68, Issue 1, pp 93-102 <http://www.healthpolicyjrn.com/article/S0168-8510%2803%2900172-6/abstract>
- Kazembe L. N., Kleinschmidt I. and Sharp B. L. (2006). Patterns of malaria-related hospital admissions and mortality among Malawian children: an example of spatial modeling of hospital register data. *Malaria Journal*. Volume 5, issue 1. pp 93.
- De La Cruz N., Benjamin C., Kirk D., Bobbi G., Natasha I., Stephen A., and Robb D. (2006). Who sleeps under bed nets in Ghana; A doer/non-doer analysis of malaria prevention behaviors. *Malaria Journal*. Volume 5, issue 1. pp 61-65.
- Y Y., Elizabeth K.M., John K., and Frederick M. (2007). Assessing the risk of self-diagnosed malaria in urban informal settlements of Nairobi using self-reported morbidity survey. *Malaria Journal*. Volume 6, issue 1. pp 71-76.
- Riedel N., Penelope V. J. M. M., Laura G., Elizabeth C. K., Victor M. and Rick W. S. (2010). Geographical patterns and predictors of malaria risk in Zambia: Bayesian geo-statistical modeling of the 2006 Zambia national malaria indicator survey. *Malaria Journal*. Volume 9. pp 37.
- Daliana P. B., Mlot C., Ferreira F. L. and Vincent J. L. (2003). Infection Probability Score (IPS) as a method to help assess the probability of infection in critically ill patients. *Critical Care Medicine*. Volume 31, Issue 11. pp 2579-2584.
- Reid H., Haque U., Clements A. C. A., Tatem A., Vallely J. A., Ahmed S. M., Islam A., and Haque R. (2010). Mapping Malaria Risk in Bangladesh Us-

ing Bayesian Geostatistical Models. *American Journal of Tropical Medicine and Hygiene*. Volume 83, issue 4. pp. 861-867.

Essuman V. A., Christine T. N. A., Birgitte S. A., George O. A., Jorgen A. K., Thomas A. N., and Bamenla G. (2010). Retinopathy in severe malaria in Ghanaian children - overlap between fundus changes in cerebral and non-cerebral malaria. *Malaria Journal*. Volume 9, issue 1. pp 232.

Noor, A. M., Amin, A. A., Akhwale, W. S. & Snow, R. W. 2007. Increasing coverage and decreasing inequity in insecticide-treated bed net use among rural Kenyan children. *PLoS Med*, 4, e255.

Chuma, J., A. Amin, A. Nyandigisi, and G. Tetteh. 2008. Access to prompt and effective malaria treatment in Kenya: a review of the literature. Submitted to the U.S. Agency for International Development by the Strengthening Pharmaceutical Systems (SPS) Program. Arlington, VA: Management Sciences for Health.

Ye Y., Elizabeth K.M., John K., and Frederick M. (2007). Assessing the risk of self-diagnosed malaria in urban informal settlements of Nairobi using self-reported morbidity survey. *Malaria Journal*. Volume 6, issue 1. pp 71-76.

Joseph Okebe, Julia M., Eugene L K., Serign J.C., Fanta N., Simon C. and Kalifa B. (2014). A comparative case control study of the determinants of clinical malaria in the Gambia. <https://doi.org/10.1186/1475-2875-13-306>

Adazu K, Lindblade KA, Rosen DH, Odhiambo PO, Kwach J, Van Eijk AM et al. Health and demographic surveillance in rural western Kenya: a platform for evaluating interventions to reduce morbidity and mortality from infectious diseases. *The American Journal of Tropical Medicine and Hygiene*. 2005; 73(6): 1151-1158

Abdur Razzaque Sarker, Marufa Sultana, Nurnabi Sheikh, Rashidul Alam Mahumud, Tania Jahir and Ziaul Islam (2017). Prevalence and associated determinants of malaria parasites among Kenyan children. <http://dx.doi.org/10.1186/s41182-017-0066-5>