University of Nairobi

School of Mathematics

ISSN: 2410-1397

Master Project in Biometry

Trends in Low Birthweight Deliveries as an Indicator of Malaria Transmission

Research Report in Mathematics, Number 52, 2020

Edwin Kiplagat Kiplelgo

November 2020

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Biometry

Trends in Low Birthweight Deliveries as an Indicator of Malaria Transmission

Research Report in Mathematics, Number 52, 2020

Edwin Kiplagat Kiplelgo

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

Master Thesis
Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Biometry

Submitted to: The Graduate School, University of Nairobi, Kenya

Abstract

Low birthweight (LBW < 2500 g) is a phenomenon that is more pronounced in developing countries where infectious diseases are most prevalent. Malaria infection in all endemic areas of Sub-Saharan Africa has become an important factor associated with LBW during pregnancy with increased susceptibility to mothers of lower parities. Hence LBW can serve as an indicator of malaria transmission. In Kenya stable malaria occur in most parts of the coast and the western regions. Part of its control mechanism is to study the traits that are linked to its spread. This study examined trends of LBW prevalence in Kilifi Health Demographic and Surveillance System (KHDSS) area. Data on birth deliveries from Kilifi County Hospital (2006-2019) were used to study trends over time. Trend significance was assessed using the Mann-Kendall test while variations of LBW prevalence were assessed by the monthly seasonal indices obtained from the Moving Average Method. Change point analysis was conducted to establish point in time when significant change in LBW prevalence occurred. Seasonal Autoregressive Integrated Moving Average model that described the LBW prevalence over time was fitted to assess the trend in the predicted values. Additive Logistic Regression was used to obtain Odds Ratio of LBW among primiparity with reference to multiparity and interpreted in relation to the contextual information regarding the changing landscape of malaria transmission. Spatial Scan Statistic (SaTScan) was used to identify local clusters of low birthweights. Findings from the study revealed a significant decreasing trend of LBW prevalence during the study period. Higher prevalence rates were observed in the southern part of KHDSS depicting high malaria transmission as compared to the Northern region. Results from the change point analysis indicated a significant change in LBW prevalence at around 2014. Variations of LBW prevalence could be explained by changes in the climatic conditions, with increased prevalence experienced shortly during the rainy periods. LBW clusters were identified in various parts of the KHDSS. Odds ratios for LBW among the primiparity could be used to define the transition of malaria in Kenya. Findings hereby, can help the government improve on the measures to combat malaria transmission in the mostly affected areas.

Keywords: Low birthweight, Malaria, change point, time series, spatial distribution.

Master Thesis 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 dissertation 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

Edwin Kiplagat Kiplelgo
Reg No. I56/24857/2019

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

Signature Date

Dr. Nelson Owuor
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: onyango@uonbi.ac.ke

Signature Date

Dr. Amek Nyaguara
KEMRI-Wellcome Trust,
Box 230, 80108 Kilifi, Kenya.
E-mail: ANyaguara@kemri-wellcome.org

Dedication

I dedicate this project to my caring parents and siblings, who frequently checked up on me during my period of research.

Contents

List of Figures

List of Tables

Abbreviations.

LBW - Low Birthweight

WHO - World Health Organization.

ANC - Antenatal Care.

KCH - Kilifi County Hospital.

ITN - Insecticide treated Nets.

LLIN - Long lasting Insecticide Nets.

IPTp - Intermittent Preventive Treatment during pregnancy

SP - Sulfadoxine pyrimethamine

KHDSS - Kilifi Health and Demographic Surveillance System.

CPA - Change-point Analysis.

SARIMA - Seasonal Autoregressive Integrated Moving Average.

ARIMA - Autoregresive Moving Average Integrated Moving Average.

AIC - Akaike Information Criteria.

BIC - Bayesian Information Criteria.

IQR - Interquartile Range.

MK - Mann-Kendall.

ACF - Autocorrelation Function.

PACF - Partial Autocorrelation Function.

Acknowledgments

A deep gratitude to the Almighty God for providing an enabling environment for the completion of this research. This work was supported through the DELTAS Africa Initiative Grant No. 107754/Z/15/Z-DELTAS Africa SSACAB and additional support from DELTAS Africa Initiative [DEL-15-003]. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government.

My ultimate appreciation goes to Dr. Amek Nyaguara, Prof. Bob Snow, Prof. Philip Bejon and Alice Kamau, for their relentless guidance on the progress of this work. Further acknowledgment goes to IDeAL, Kilifi County hospital through KEMRI-Wellcome Trust for their great support throughout this project which included the provision of the data for analysis.

A special thanks to Dr. Nelson Owuor who willingly provided guidance on statistical methods during this research.

Lastly, I wish to acknowledge the University of Nairobi and more so to the School of Mathematics for according me a smooth learning period.

Edwin Kiplagat Kiplelgo

Nairobi, 2020.

1 INTRODUCTION

1.1 Backgroud

Malaria infection still remains a public health challenge around the globe. It is an infection caused by plasmodium parasites transmitted to humans through bites of infected female Anopheles mosquitoes. According to the World Health Organization (WHO), a worldwide total of 228 million malaria cases and 405,000 deaths in 2018 were due to malaria infection, WHO (2019). The most vulnerable groups impacted by malaria are expectant mothers and children below the age of five.

Pregnant women are prone to infection with malaria Desai et al. (2007) and this risk is attributed to the immunological changes in pregnancy and to the unique predilection of a group of *plasmodium falciparum* parasites to sequester in maternal blood spaces of the placenta, (Rogerson et al., 2007). Malaria in pregnancy poses a negative pregnancy outcome to both the mother and the developing fetus. As a result, pregnant women having the infection are more likely to deliver low birthweight (LBW) infants. Walker et al. (2014).

Low birthweight (LBW) has been described by WHO as a weight at birth that is less than 2.5 Kg irrespective of gestational age. This empirical cut-off for international comparison was focused on epidemiological findings that infants weighing about 2.5 Kg were about 20 times more likely to die than normal weight babies, (Kramer, 1987). LBW is an important factor associated with greater risk of illness, increased vulnerability to childhood disease, reduced odds of infant survival, long-term physical and mental deficits. (Metgud et al., 2012) . Mmbando et al. (2008), also affirmed that children having low birthweight are at high risk of morbidity, impaired cognitive development and growth retardation.

LBW is a phenomenon which is more pronounced in developing countries where infectious diseases are extreme. Statistics from WHO in 2018 indicated a global estimate of 15.5% LBW prevalence which amounted to about 20 million low birthweight babies each year, with developing countries contributing to about 96.5%.

Many studies have attributed LBW to malaria infection during pregnancy. Shulman et al. (2001) reported that past placental malaria was associated to LBW;Unger et al. (2019) in their study reported that peripheral microscopic *p. falciparum* at the time of delivery was associated with low birthweight.

Interventions aimed at minimizing the risk of malaria infection during pregnancy have shown improvement in the outcomes of birthweight among primigravid women. A meta-analysis of these findings from randomized control trials suggests that primigravidae under an intervention against malaria infection deliver infants that are 101 g heavier on average compared to those who are not subjected to any intervention, (Gülmezoglu and Garner, 1998). Several other studies have gone an extra mile and reported that the occurrence of low birth weight due to malaria infection declines as the gravidity increases, (Mutabingwa et al., 2005) implying that the higher risk among the primigravids could be used to study malaria transmission (Brabin et al., 1999). Thus, this retrospective study focuses on evaluating trends of low birthweights and to assess the endemicity of malaria using LBW as an indicator of transmission.

1.2 Statement of the problem.

Normal Weight at birth implies increased likelihood of infant survival. To a greater extent, low birth weight on the other hand remain to be a significant risk factor to infant and child mortality. It may also contribute to developmental complications in later stages of adulthood. Studying the trends of LBW in relation to its causing factors help to identify the most affected areas and to effectively plan to alleviate or improve on preventive mechanisms against the leading causes of low birth weight.

Malaria infection during pregnancy is an important factor which poses adverse effects to the mother and the child. The infection may lead to intrauterine growth restriction and preterm birth, (Desai et al., 2007) which are the main contributing factors to LBW. It has been estimated in Sub-Saharan Africa that approximately 900,000 LBW deliveries related to placental infection due to *plasmodium falciparum* would occur yearly if MIP preventive measures are not implemented, (Walker et al., 2014). In Kenya, Low birthtweight has been estimated to be 11.5%, Blencowe et al. (2019).

The infection however, has been documented to be different in parity with those of lower parity having increased susceptibility to malaria infection, BRABIN (1991); Rogerson et al. (2007). This finding postulate that the excess risk of LBW in first pregnancies can serve as an indicator of malaria transmission and exposure, (Brabin et al., 1999). Many studies in kenya have focused on the epidemiology of malaria but none has used LBW has an indicator of malaria transmission. We therefore use the low birthweight in both the primigravid and multigravid women as a metric to explain the dynamics of malaria transmission and exposure.

1.3 Objectives

1.3.1 Main objective.

The main objective of the study was to evaluate trends in low birthweight deliveries to explain malaria transmission and exposure.

Specific objectives.

- To evaluate the trend and seasonal variations in LBW prevalence.
- To evaluate the endemicity of Malaria in the study region.
- To establish the spatial distribution of low birthweights cases.

1.4 Significance of the study

Significant goal in any health related research is to gain insight about a disease or health condition under consideration with the aim of developing health interventions. Studying malaria dynamics in relation to low birthweight is a simple and an inexpensive approach that can be implemented easily. At the peak of this research, we would have assessed the evolving patterns of low birthweight over time. Moreover, malaria transmission in relation to low birthweight would be known. The researchers expect that the findings derived from this work will play a significant role in alleviating two aspects in the health spectrum: Knowledge on the transmission of malaria will be known hence improving on measures to alleviate its spread and in the long run other malaria-related conditions like low birthweight will be controlled.

1.5 Structure of the thesis.

This work comprises five chapters. The first chapter presents an overview for the all project and it is made up of the following sections; introduction, problem statement, objectives and significance of the study. The second chapter presents a review of previous work on low birthweight, Malaria in pregnancy, endemicity of malaria and LbW and finally on strategies to prevent malaria in pregnancy. Chapter three describes the study design, study area and the population. Moreover, a detailed explanation on the methodologies that have been used to analyze the data is provided; these included Mann-Kendall test for trend, change-point analysis, time series models and model building procedure using the Box-Jenkins approach, evaluation of malaria transmission and lastly, spatial pattern analysis to study clustering of LBW cases has been discussed. Chapter four gives

the descriptive statistics and analysis of the data using the aforementioned methodologies. The last chapter discusses findings, conclusions, limitations, recommendations and suggestions for future studies.

2 LITERATURE REVIEW

2.1 Overview

This chapter presents works that have been addressed by other authors on LBW, malaria during pregnancy, endemicity and ways to prevent malaria during pregnancy.

2.2 Low birthweight

The WHO describes LBW as birthweight below 2500 g irrespective of gestational age and typically refers only to live births. Birth weight is determined within the first hour of life prior to significant postnatal weight loss, Blanc and Wardlaw (2005). LBW in infants is a product of preterm birth (PTB), Intrauterine growth restriction (IUGR) or both (Kramer, 1987). A birth is deemed preterm if it takes place before the 37th week of pregnancy.

LBW is still a global public health problem and is associated with a range of short and long term effects. It is a factor associated with higher risk of illness, increased vulnerability to childhood diseases, reduced odds of infant survival, long-term physical and mental deficits, (Metgud et al., 2012). Mother-related causes to the LBW range from nutritional status, tobacco and alcohol use, non-communicable diseases and infections including malaria, Wardlaw (2004).

WHO's goal has been to realize a 30% decrease in the number of babies born below 2.5 Kg by 2025. Recently,in 2018 it was estimated that 15% to 20% of all births around the world were of LBW accounting for 20 million births annually.

2.3 Malaria

Malaria stands out to be one among the world's major cause of morbidity and mortality. It continues to be a significant life-threatening disease despite the immense measures that have been set to control its spread. Infection in humans is caused by the Plasmodium parasites which are transmitted through the bites of infected female Anopheles mosquitoes, (Autino et al., 2012). Malaria-transmitting plasmodium parasites include; *p. falciparum, p. malariae, p. Ovale* and *p. vivax*

In Africa, *p. falciparum* is predominant and it accounts for the adverse effects of the disease. In 2018, it accounted for 99.7% of all the reported malaria cases in WHO African region, (WHO, World malaria report 2019). Symptoms of Malaria according to WHO can

develop between 8 and 25 days after the infection; they include headaches, fever, vomiting, shivering, joint pain, abdominal pain, convulsions, nausea or muscle pain. Severe malaria can result into having seizures, abnormal postures or even coma. The most susceptible groups are young children under the age of five whose immune system is yet to develop, pregnant women with weakened immune system due to pregnancy and travelers traveling from areas with little to no transmission to high transmission areas.

World Health Organization in their recent report estimated that a worldwide total of 231 million cases, [95 % CI; 211-259 million] and 416,000 deaths in 2017 were due to malaria infection. In 2018, there was a slight decrease in the number of deaths reported (405,000) while the number of cases decreased to 228 million, [95 % CI; 206-258 million]. The WHO African Region had most of the reported cases in 2018 (213 million), equivalent to 93 % of the total cases.

Like any other part of Sub-Saharan Africa, malaria in Kenya is a serious public health issue giving rise to morbidity and mortality. The temporal presentation of malaria in the country since 1990 as described by (Macharia et al., 2018) varied with the changing landscape of disease management, climate anomalies and vector control. From 1990 to 2015, the national mean of *plasmodium falciparum* parasite prevalence (*Pf* PR) declined by 87.7%. For the period of 1990 and 1998 the mean value of (*Pf* PR) had remained constant at 21.2% until when a small decrease was reported in 1998 and 1999 after which it remained constant until 2003. Between 2003 and 2007 a greater decrease of 81% (from 17.1% to 3.2%) in the national mean of (*Pf* PR) was recorded. The same level was maintained from 2007 until when a rise was seen from 2011 to 2014 and eventually fell to 2.6% in 2015.

2.4 Malaria during pregnancy.

Pregnant women are vulnerable to infections from malaria, Desai et al. (2007). The vulnerability arises when a womans immmune system is weighed down by the pregnancy thus rendering her more prone to the infection. In Sub-Saharan Africa, Malaria during pregnancy still poses substantial adverse effects on mother's health and the developing fetus, (Walker et al., 2014). It causes intrauterine growth restriction, preterm delivery and maternal anemia,(Desai et al., 2007) which may lead to LBW. The Intrauterine development is impaired when erythrocytes infected with Plasmodium parasites-sequester in the placental intervillous space disturbing the flow of nutrients to the fetus and creates a reservoir of inflammation, (Rogerson et al., 2007). The high density infection of parasites according to Guyatt and Snow (2004) may also contribute to the consumption of oxygen and glucose meant for the development of the fetus, hence the inflamatory reaction due to the infection induces immunotollerance in the offspring as suggested by, Cot et al. (2003). In a study done by Cottrell et al. (2007) found that placental infection and peripheral infection during pregnancy associated significantly with low birthweight in infants. But most notably is the moment when malaria infection has adverse effects on the fe-

tal growth during pregnancy. Cottrell in their research took into account three periods of pregnancy; less than 4 months of pregnancy, 4-6 months and more than 6 months of pregnancy. They found that peripheral infection was correlated significantly with LBW, however, the infection was only significant at the end of the pregnancy upon adjustment of the cofactors. Moreover, a cohort study on the timing and effect of malaria during pregnancy on child growth and morbidity conducted in Uganda (De Beaudrap et al., 2016) showed that infection with malaria occurring 12 weeks before birth posed a higher risk of height and weight growth retardation. Findings above show that malaria infection occurring during the last trimester contributes significantly infants' LBW.

The prevalence and the effect of malaria infection during pregnancy can differ with maternal gravidity, (Mutabingwa et al., 2005) This is because women develop adaptive immunity which reduces the risk of placental infection with succeeding infected pregnancies, (Rogerson et al., 2007). Study conducted by Okiring et al. (2019) in highly endemic area indicated that, gravidity was the most significant risk factor for microscopic parasitemia, with primigravid women having almost double the risk when compared to multigravida.

2.5 Endemicity of Malaria and LBW.

Malaria endemic areas can be categorized into high and low-transmission regions (Beier et al., 1999). Most areas of Sub-Saharan Africa have been described as having a high transmission of malaria, while in some other areas the transmission is moderate to low. In all endemic regions, pregnant women are at risk of malaria infection. Study by Brabin (1983) reported that in the last trimester malaria prevalence was highest irrespective of the area. Moreover, in a subsequent study by Brabin and Rogerson (2001), found that pregnant women were more vulnerable to malaria infection than either before or after pregnancy owing to their decreased immunity due to pregnancy and seasonal fluctuations in malaria transmission.

However, despite the fact that pregnant women are at risk to infections in all endemic regions, they are asymptomatic in areas characterized by high malaria transmission, (Nosten et al., 1991) compared to areas of low transmission where Plasmodium falciparum infection is symptomatic with typical symptoms of fever, vomiting, headache, nausea among others.

While the prevalence of malaria parasites is lower with rising gravidity, the infection affects all parity classes in low-transmission areas, (Nosten et al., 1991). whereas in areas of high malaria transmission primigravidae are at higher risk of infection and its negative consequences compared to multigravidae, Steketee et al. (1996).

With a greater link of LBW to malaria during pregnancy, maternal malaria in all endemic areas is related to low birth weight, (Menendez, 1995). Where malaria transmission is

high, the risk of low birthweight was approximated to be double if women had placental malaria, (Guyatt and Snow, 2004) with primigravidae experiencing greater impact.

2.6 Anemia and malaria

Anemia is a condition that is attributable to the concentration of hemoglobin falling below normal. It is a major public health challenge affecting mostly pregnant women and young children. Anemia during pregnancy according to the WHO is characterized as hemoglobin (Hb) levels that is less than 11.0 g/dL. Concentrations that fall below a specified cut-off point impairs the blood's capability to transport oxygen across the body.

Anemia during pregnancy is associated with adverse maternal and neonatal health effects such as, but not limited to, intrauterine growth retardation, preterm births, LBW and maternal mortality, (Ouma et al., 2007). Study done by Desai et al. (2007) further alludes that high transmission of malaria leads to maternal anemia which is a significant factor towards LBW in infants resulting due to limitations on the development of the fetus or premature delivery.

In pregnancy, the causes of anemia are multifactorial. Among the causes of anaemia is infection by malaria which may occur through excess removal of non-parasitized erythrocytes (Ekvall, 2003) thus anaemia stands to be the most common symptom of malaria in pregnancy. A study by Cottrell et al. (2015) reported that infection by Submicroscopic Plasmodium parasites was associated with decreased mean levels of Hb during pregnancy and the effect being more pronounced in primigravidae compared to multigravidae.

Decreased Hb concentrations increases the chances of LBW Bahizire et al. (2018). This mostly occur in the case of severe anaemia (Hb < 7 g/dl) which is a risk factor for intrauterine growth restriction. During pregnancy, anaemia causes chronic hypoxia or oxidative stress which stimulates the production of corticotropin releasing hormone that increases cortisol production which inhibits fetal growths.

2.7 Malaria in pregnancy: The situation in Kenya

In its bid to combat malaria during pregnancy, Kenya adopted the three-pronged approach of the WHO which includes Intermittent preventive treatment in pregnancy (IPTp), distribution and use of nets treated with insecticides and effective case management, (Robert et al., 2004). The provision of IPTp was first introduced in Kenya in 1998 upon which at least two doses of sulphadoxine pyrimethamine were given during the second and third trimesters to all pregnant women. The policy was later updated in 2009 in line with the WHO recommendations to restrict IPTp to women living in endemic areas. Currently, IPTp policy recommends treatment of malaria in malarias endemic counties like the Coast and Western Kenya. The treatment involves prescribing sulfadoxine-

pyrimethamine (SP) to pregnant women via directly observed therapy at each antenatal care beginning in the second trimester with corresponding doses at least 4 weeks apart.

Kenya Malaria indicator Survey (KMIS, 2015) was conducted on women who had live births 2 years before the survey to determine the utilization of IPTp. The findings indicated that nationally 51% of women received at least one dose of IPTp, 35% received at least two doses and 22% got at least three doses. Taking endemic areas into consideration, 77% of women received at least one dose, 56 % received two or more doses and 38% received the 3 or more doses which is currently recommended. The survey reported an improvement in IPTp coverage, but still too small was the 38% who received the recommended dose.

Using insecticide-treated bed nets as a preventive mechanism against mosquito bites serves as an important measure to control malaria during pregnancy. The survey assessed ownership of mosquito nets and the outcome was that 63% of households possessed at least one long-lasting insecticidal net (LLIN) while 37% owned more than 1. That averaged to 1.3 LLINs per household which was an increment from the previous 0.8. From the survey, 40% of the households in Kenya had hit a universal LLIN coverage of one net per two people sleeping in the night prior to the survey. The study also reported that the night prior to the survey, 58% of pregnant women (aged 15-49) had slept under LLIN.

2.8 Strategies to prevent malaria in pregnancy

Despite the detrimental effects posed by malaria during pregnancy, it is an infection which can be prevented. Measures to combat its adverse effects have been suggested and even implemented. The WHO recommends a three-pronged approach to malaria infection prevention and management in pregnant women in Sub-Saharan Africa. These strategies are discussed in the next subsections.

2.8.1 Insecticide treated nets

Sleeping under Insecticide treated Nets (ITNs) remains to be a core method in the protection against mosquitoes transmitting malaria. Use of ITNs in high transmission areas have exhibited profound usefulness, (Rogerson, 2017). As evidenced in reviews from Africa, ITNs could be protective of malaria and decreased LBW by 29 %, reduced miscarriages and still births in paucigravidae by 33 % and decreased placental parasitemia in all gravidae by 23 %, (Eisele et al., 2010). Study done by Bahizire et al. (2018) showed that having slept under ITN the night before the antenatal visit was protective against LBW.

Given the implementation of this strategy by WHO, the proportion of pregnant women sleeping under ITNs was still low. This is according to the World malaria report of 2019 by

WHO where reports in 2018 showed that, 50% of Sub-Saharan Africa's population slept under the ITN.

Bednet distribution in Kilifi, Kenya, had long begun in the 1990s with an estimated usage of less than 6% Snow et al. (1992). By the year 2005, insecticide treated nets had been made available to pregnant mothers and children. Acquisition was at the maternal and child health clinics. A free ITN mass distribution drive carried out by the then Kilifi District Management Team later in September 2006 led to an expanded coverage across kilifi from 0.25 to 0.5 nets per person, (O'Meara et al., 2008). Provision of nets to the community later after the campaigns were made available at 50 Kenyan shillings. This scenario changed when a new regulation was adopted in January 2009 which saw the pregnant women and children under the age of five receive bednets freely whilst the rest of the population acquired at a cost of 50 Kenya shillings. A further mass distribution of bednets was carried out in July 2012 and October 2015 as a universal coverage for people at risk of malaria. The government also initiated another mass delivery of long-term insecticide nets in 2017, which ended in March 2018.

2.8.2 Intermittent preventive treatment in pregnancy (IPTp)

In Africa, the WHO recommends prevention of malaria in pregant women in regions experiencing moderate to high transmission to receive IPTp with with sulfadoxine-pyrimethamine administered during the antenatal care visits. The drug administration during antenatal visits is done under supervision, (Peters et al., 2007) whether or not the patient is infected. IPTp reduces placental parasitemia, maternal anaemia, LBW and neonatal mortality.

In 2016, the WHO released a new guideline that pregnant women in endemic areas be provided with at least three doses of IPTp-SP beginning at their second trimester, with at least one month after each dose. This guideline from WHO has been proven to be effective by studies. A multisite study found that, despite the growing resistance to SP which is common in eastern and southern Africa, IPTp is still effective against LBW, (Desai et al., 2016). In a Meta-analysis Study by Kayentao et al. (2013) revealed that at least three doses of SP were associated with 23% reduced risk of low birthweight among HIV-negative women relative to two doses.

But still the percentage of women receiving IPTp in Sub-Saharan Africa is highly variable. In 2018 it ranged between 3% in Namibia to 58% in Burkina Faso. During the same time, 20% percent on average of eligible women received at least three doses of IPTp.

Information from the Demographic Healthy Surveys-DHS in Kenya, indicate that the proportions of women receiving IPTp during the ANC visits has been increasing since 2003. 4% of the women were reported to have received IPTp during the visit in a survey done in 2003. Later in 2008-2009 DHS, the proportion rose to 14%. In 2014, 17% of women

were indicated to have taken two or more doses of sulphadoxine pyrimethamine while 10% had received three doses or more.

In Kilifi county, with regard to the DHS 2014, the proportions of women who during the ANC visit received one or more doses were 65.2%, two or more doses were estimated to be 41.8% while those who received more than three doses were 28.5%.

2.8.3 Effective case management

Pregnant women in malarious areas are at increased risk of the adverse impacts initiated by the infection of malaria. Early diagnosis and treatment help lessen placental malaria leading to decreased adverse effects on the fetus, (Nosten et al., 2007). Effective case management of malaria during pregnancy calls for biological diagnosis of malaria prior to treatment. The biological diagnosis is necessary to avoid nonessential exposure to antimalarials drugs for pregnant women and their fetus, (Nosten et al., 2007). Malaria infection can be detected through microscopy examination or by use of rapid diagnostic test that can detect a particular parasite antigen. However, microscopy sensitivity is ineffective in situations where asymptomatic low parasite density sequesters in the placenta which impairs malaria diagnosis and treatment in pregnant women. A study by Cottrell et al. (2015) indicated light microscopy missed at least half of *p.falciparum* infections in maternal peripheral blood. Therefore, a more sensitive Polymerase Chain Reaction (PCR) has helped in revealing the burden of infection during pregnancy. A blood test should be provided for malaria and positive cases be handled accordingly for any pregnant woman seeking for an antenatal care in areas known to transmit malaria, and the lack of evidence of parasites in the peripheral blood does not rule out infections, (Nosten et al., 2007).

For cases of severe anemia in pregnant women from endemic areas, an effective antimalarial is administered whether or not there are peripheral parasitemia or fever records. Due to the growing resistance to both chlorine and sulfadoxinepyrimethamine, WHO recommends the use of artemisinin-based combination therapy (ACT) to pregnant women having severe anemia and living in areas with a great risk of P.falciparum in their last 2 trimesters and another effective antimalarial in the first trimester.

3 MATERIAL AND METHODS

3.1 Overview

The present chapter elucidates the structure and the methods that have been used to address the problem statement for this research. The study area and the population have been highlighted. Ways of data extraction and inclusion criteria also discussed. Finally, statistical methods have been introduced.

3.2 Study Design.

The study employed a retrospective longitudinal study design.

3.3 Study area.

Kilifi County occupies an area of $12,610\ Km^2$ with a population of 1,453,787 according to the 2019 Kenya Population and Housing Census. The eastern part of the county stretches 65 Km along the Kenyan coast line and extends 90 Km from the widest point of the coast to the inland, (O'Meara WP et al 2008). The county has seven sub counties, namely; Kilifi North, Kilifi South, Kaloleni, Rabai, Malindi Ganze and Magarini. The area is further subdivided into 36 administrative locations. Kilifi Health and Demographic Surveillance System (KHDSS) was established as a record of migration events and other vital statistics. The area under surveillance covers 891 Km^2 and stretches 35 Km^2 along the coastline, (Scott et al., 2012). The county experience a bimodal rainfall pattern with long rains occurring from April to June, while short rain falls in October and November. Kilifi County Hospital is a level 4 government health facility situated within Kilifi Township and serves as first referral centre for children and adults within the KHDSS and other parts of the County. The hospital has a maternity ward that attends to pregnant mothers seeking antennal care and delivery services. According to the latest records the department registers more than 4,000 deliveries yearly.

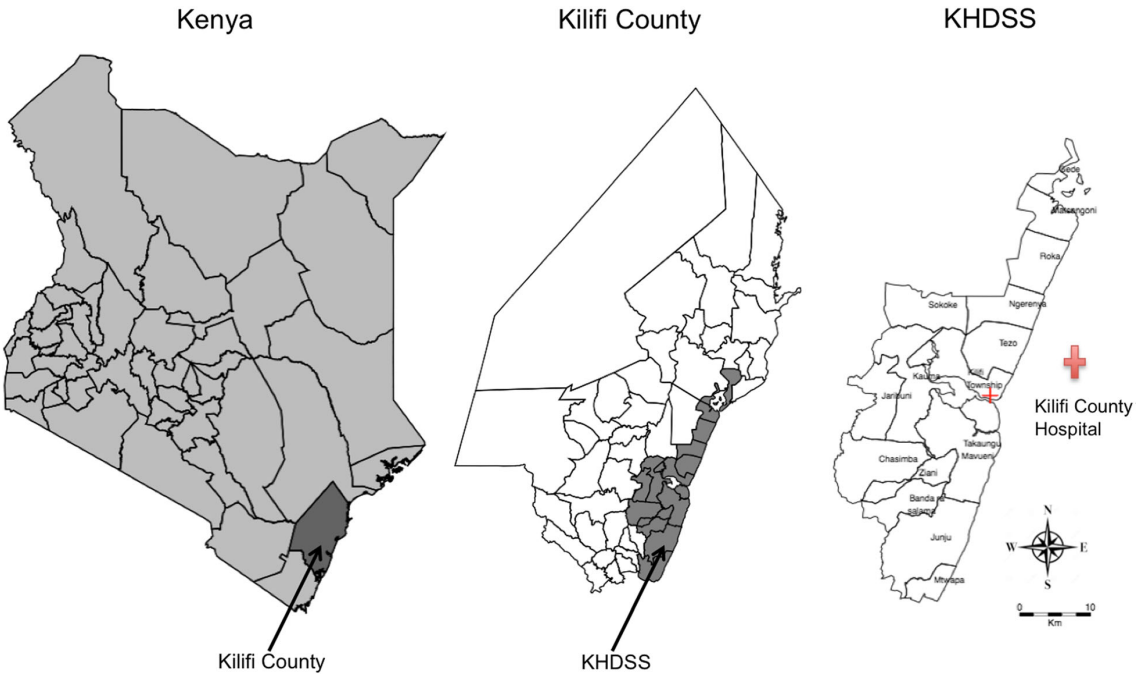

Figure 1. Map of Kenya on the left indicating the location of Kilifi county, at the center is Kilifi County map showing the location of KHDSS and on the right is locations in KHDSS

3.4 Study population

The study considered all pregnant women who sought delivery services at the Kilifi County Hospital between June 2006 and December 2019.

3.5 Ethical Approval

This study got approval from the Kenya Medical Research Institute Scientific and Ethics Review Unit (SSC 1778).

3.6 Data extraction.

Information on birth registry was obtained from two sources. Between June 2006 and December 2010 data were extracted from the hospital registries at the Kilifi County Hospital (KCH). These data were extracted using a standardized extraction form in Microsoft Excel designed to capture relevant information. The second source of information was a database hosted in the KEMRI-Wellcome Research Programme where data had been systematically recorded between January 2011 and December 2019. Maternal data collected include parity, place of residence and the history of Intermittent Preventive Treatment during pregnancy, while infant information include birth weight, sex and the delivery outcome .

3.7 Inclusion criteria.

Records used for analysis had to fulfill the following inclusion criteria: First, all live but singleton births were considered to form part of the analysis. Secondly, births that were defined within the KHDSS area were included.

3.8 Statistical data analysis.

Data extracted from the two sources were cleaned, merged and descriptive analysis obtained in STATA for windows version 14.2 (StataCorp, College Station, TX, USA) and R version 3.5.1 was used in time series analysis. Information on birth weight was transformed into a categorical variable; weights below 2500 g were categorized as LBW, otherwise normal weight. Thereafter, monthly prevalence of low birthweight were computed.

Summaries of non-normal quantitative variables were presented using median and interquartile range (IQR). Birthweight differences between groups (parity, sex) were compared using student t-test. Statistical difference of low birthweight proportions between groups (region, parity and gender) were assessed using Z test. Additive Logistic Regression was used to obtain Odds Ratio of LBW among primiparity with reference to multiparity and interpreted in relation to the contextual information regarding the changing landscape of malaria transmission.

Long term trend in prevalence of LBW was assessed using Mann-Kendall (MK) test. The magnitude of the trend was obtained using the Theil Sen method.Since MK test is significantly affected by the presence of periodicities and serial autocorrelation in a time series data, Autocorrelation function (ACF) and decomposed time series was used to check for the autocorrelation and to assess the seasonal component, respectively.

Change point analysis was conducted to establish point in time when significant changes in prevalence of LBW were experienced. A time series model (Seasonal auto regressive moving average) that described the data was fitted. To analyze the seasonal variations in the prevalence of LBW, we applied the moving averages method.

The overall and stratified LBW ratios of primiparity group in relation to multiparity were computed upon which trend analysis was evaluated. Lastly, Spatial distribution of LBW prevalence and clusters were geographically mapped.

3.9 Mann-Kendall (MK) test for trend.

Mann-Kendall is a non-parametric test commonly used to evaluate the trend of a time series data. Unlike other test like the regression coefficient test which assumes the normality of the data, Mann-Kendall test statistic is distribution free since it only relies on the ranks of the observations. It tests the null hypothesis of no monotonic trend.

H_0: No monotonic trend in the series.
H_1: Presence of monotonic trend in the series.

For a time series $Z_t = Z_1, Z_2, Z_3, ..., Z_n$, the test statistic is given by;

$$S_{MK} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} D_{ij} \tag{1}$$

Where $D_{ij} = \text{Sgn}\ \ (Z_j - Z_i)$; $j > i$ and n is the time series observations

$$\text{Sgn}\ (Z_j - Z_i) = \begin{cases} +1, & \text{if}\ \ Z_i < Z_j \\ 0, & \text{if}\ \ Z_i = Z_j \\ -1, & \text{if}\ \ Z_i > Z_j \end{cases}$$

For $n \geq 10$, the variance is given by;

$$Var(S_{MK}) = \frac{n(n-1)(2n+5)}{18} \tag{2}$$

When the series has tied ranks then;

$$V^*(S_{MK}) = \frac{n(n-1)(2n+5)}{18} - \sum_{i=1}^{m} \frac{r_i(r_i-1)(2r_i+5)}{18} \tag{3}$$

Where m represents the number of clusters of tied ranks while r_i is the total observations in a given cluster.

MK test statistic tends to normal distribution as the size of n gets larger. Significance of the trend can thus be tested by comparing the computed Z_s statistic with the standard normal value. For this analysis we chose α=0.05.

$$Z_s = \begin{cases} \frac{S_{SK}-1}{\sqrt{Var(S_{SK})}}, & \text{if}\ \ S_{SK} > 0 \\ 0, & \text{if}\ \ S_{SK} = 0 \\ \frac{S_{SK}+1}{\sqrt{Var(S_{SK})}}, & \text{if}\ \ S_{SK} < 0 \end{cases} \tag{4}$$

Where Z_s is the standardized value. An upward trend is represented by positive values of Zs while negative values represent a downward trend in the time series.

MK test gives the the direction and the significance of a trend but does not provide the magnitude of the slope. We estimated the magnitude by using the Sen's Slope method (Sen, 1968). To illustrate this concept; assume a sample of M pairs of observations,

$$Q_k = \frac{Z_j - Z_i}{j - i} \text{ for } k = 1, 2, \ldots, M \tag{5}$$

Where;
Z_j and Z_i are time series observations at times j and i respectively for $j > i$

The resulting M values of Q_k are ranked in ascending order. The Sen's Slope is then estimated as;

$$Q_{Med} = \begin{cases} Q_{\frac{M+1}{2}}; & \text{if } M \text{ is odd} \\ \frac{Q_{\frac{M}{2}} + Q_{\frac{M+2}{2}}}{2}; & \text{if } M \text{ is even} \end{cases} \tag{6}$$

The value of Q_M indicates the magnitude of the slope. while the positive and negative signs of of Q_M indicate whether the trend is increasing or decreasing respectively. Evaluation of slope's significance was at $\alpha = 0.05$.

3.10 Change point analysis (CPA)

A change point is an instance in time when the time series properties such as mean or variance before and after a certain time point differ. Change Point Analysis in a time series data aims at detecting and localizing points at which these statistical properties change. Assuming a time series $Z_t = (Z_1, Z_2, Z_3, \ldots, Z_n)$, a change point within the set of the series occur when a time point $\tau \in (1, 2, \ldots, n-1)$ exist such that the statistical properties of $Z_1, Z_2, \ldots, Z_\tau$ and $Z_{\tau+1}, Z_{\tau+2}, \ldots, Z_n$ differ.

3.10.1 Change point detection.

A change point in the data can be evaluated by applying the general likelihood Ratio (LR) method. LR is commonly used in assessing the goodness of fit of two nested models. In this context, the LR test is used to compare two models, one with no changepoint and another having a changepoint. More precisely, to test for a change point in mean then the hypotheses to test is as follows;

$H_0 : \mu_1 = \mu_2 = \ldots = \mu_n$

$H_1 : \mu_1 = \cdots = \mu_\tau \neq \mu_{\tau+1} = \ldots = \mu_n$

The maximum log-likelihood is then obtained for both hypotheses.

Under the null hypothesis, we have;

Likelihood of H_0 :

$$L(H_0) = \prod_{i=1}^{n} Pr(Z_{1:n}|\widehat{\theta})$$

Maximum log likelihood of H_0 :

$$\sum_{i=1}^{n} \log(Pr(Z_{1:n}|\widehat{\theta})$$

Where $\widehat{\theta}$ is the maximum likelihood estimate (MLE) of the parameters (for instance, the mean) and $Pr(Z_{1:n}|\widehat{\theta})$ is the pdf associated with the distribution of $(Z_{1:n})$. Generally, assuming Gaussian distribution of our data, then the likelihood function is given as;

$$L(H_0) = \prod_{t=1}^{n} \frac{1}{\left(\sqrt{(2\pi\sigma^2}\right)^n} \exp -\frac{1}{2\sigma^2}(Z_t - \mu)^2 \tag{7}$$

The maximum likelihood estimators for μ and σ^2 are;

$$\widehat{\mu} = \frac{1}{n}\sum_{t=1}^{n} z_t \quad \text{and} \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}(z_t - \bar{z})^2 \quad \text{respectively.}$$

Under the Alternative hypothesis

The alternative hypothesis represents a model with a change point at τ where $\tau \in 1, 2, 3, \ldots, (n-1)$. The likelihood function is thus;

$$L(H_1) = \prod_{t=1}^{\tau} Pr(Z_{1:\tau}|\widehat{\theta}_1) \prod_{t=\tau+1}^{i=n} Pr(Z_{(\tau+1):n}|\widehat{\theta}_2) \tag{8}$$

While the corresponding maximum log-likelihood;

$$\sum_{t=1}^{\tau} \log \Pr(Z_{1:\tau}|\widehat{\theta}_1) + \sum_{t=\tau+1}^{n} \log Pr(Z_{(\tau+1):n}|\widehat{\theta}_2)$$

More generally;

$$L(H_1) = \prod_{t=1}^{\tau} \frac{1}{\sqrt{(2\pi\sigma^2)}^n} \exp -\frac{1}{2\sigma^2}(Z_t - \mu_1)^2 \prod_{t=\tau+1}^{n} \frac{1}{\left(\sqrt{(2\pi\sigma^2)}^n\right.} \exp -\frac{1}{2\sigma^2}(Z_t - \mu_2)^2 \tag{9}$$

The maximum likelihood estimators for μ_1, μ_2, and σ^2 are respectively;

$$\widehat{\mu}_1 = \frac{1}{\tau}\sum_{t=1}^{\tau} Z_t, \quad \widehat{\mu}_2 = \frac{1}{n-\tau}\sum_{t=\tau+1}^{n} Z_t,$$

and

$$\widehat{\sigma}^2 = \frac{1}{n}\left[\sum_{t=1}^{\tau}(Z_t - \bar{Z}_\tau)^2 + \sum_{t=\tau+1}^{n}(Z_t - \bar{Z}_{n-\tau})^2\right]$$

Maximum log likeligood:

$$ML(\tau) = \log Pr(Z_{1:\tau}|\widehat{\theta}_1) + \log Pr(Z_{(\tau+1):n}|\widehat{\theta}_2)$$

The likelihood Ratio Test:

$$LR = \log\left(\frac{L(H_1)}{L(H_0)}\right) = \sum_{t=1}^{\tau} \log Pr(Z_{1:\tau}|\widehat{\theta}_1) + \sum_{t=\tau+1}^{n} \log Pr(Z_{(\tau+1):n}|\widehat{\theta}_2) - \sum_{t=1}^{n} \log Pr(Z_{1:n}|\widehat{\theta}_0)$$

The maximum of Likelihood Ratio test

$$\gamma = 2\left[\begin{matrix} Max\ LR \\ \tau \end{matrix}\right]$$

Change point is inferred if $\gamma > c$ and the MLE is the value τ that maximizes the LR. The value of c is a threshold set to test the null hypothesis. More often, c is such that $\alpha = Pr(\gamma > c|H_0)$ for a chosen level of α.

$$\widehat{\tau} = \left[\begin{matrix} argMax \\ \tau \end{matrix}\quad LR\right]$$

3.11 Concept of time series.

A time series $Z_t = Z_1, Z_2, \ldots, Z_n$ is a collection of measurements on a quantitative variable over time. Time series data produced regularly (for example daily, weekly or monthly) forms a regular time series. The analysis of data from a time series helps to draw statistics and data-related features so as to explain the driving factors and structure behind the observed data.

3.11.1 Time series components.

A time series can be decomposed into four components each with a particular pattern.

Trend component.

The trend is the time series long-term pattern. It can exhibit a positive or negative trend depending on whether the time series has an increasing or decreasing long term pattern. Trend in a time series can either be linear or nonlinear. A nonlinear trend can assume an exponential or a quadratic form.

Seasonal variations.

Seasonality takes place when the time series consistently fluctuates in the same months each year. Seasonal variations are mostly due to climate and weather conditions, or even customs.

Cyclical fluctuation:

Can be described as an up and down movement around a particular trend. The fluctuation is unforeseeable as it does not assume a constant time interval. The difference between seasonal and cyclical variation is that the latter has variable lengths and usually lasts for longer periods while the former reoccurs at regular time intervals with shorter periods less than one year.

Irregular variation.

It is the time series random component. It is the variation which can not be attributed trend,seasonal or cyclical variations. The variation is random, unforeseeable and unpredictable. It can be caused by incidents like floods or earthquakes.

3.11.2 Time series decomposition.

Decomposition entails the splitting of a time series into its various patterns which include trend, seasonality, cyclicality and the noise. Thus, time series is a function of four components.

$$Z_t = f(T_t, S_t, C_t, \omega_t)$$

Where Z_t is the time series values at different time points; T_t the secular trend; S_t the seasonal part; C_t; cyclic component and ω_t the white noise. The above function can exhibit two forms

- **Additive model.** Under additive model the assumption is that the four components are independent of one another, hence the time series is the sum of its components.

$$Z_t = T_t + S_t + C_t + \omega_t$$

- **Multiplicative model.** The assumption is that the time series components can affect one another hence not necessarily independent.

$$Z_t = T_t \cdot S_t \cdot C_t \cdot \omega_t$$

  Applying the log transformation to the multiplicative model converts to additive model.

$$\begin{aligned} \log(Z_t) &= \log(T_t \cdot S_t \cdot C_t \cdot \omega_t) \\ &= \log(T_t) + \log(S_t) + \log(C_t) + \log(\omega_t) \end{aligned}$$

3.11.3 Differencing

Differencing is a technique that helps in stabilizing a non-stationary time series. A first order difference involves computing the differences between two consecutive observations. For instance a time series Z_t can be transformed into a new set of observations having (t-1) values; ${Z^*}_t = Z_t - Z_{t-1} = \nabla Z_t$, where ∇ is the difference operator. Differencing of a time series can be done many times until stationarity is achieved.
A back shift operator which is defined by $BZ_t = Z_{t-1}$ is usually introduced for convenience. Where BZ_t represents a time series observation at lag 1. This follows that $B^2 Z_t = B(BZ_t) = B(Z_{t-1}) = Z_{t-2}$ and thus $B^k Z_t = Z_{t-k}$

A first order difference will take the form;

$$\nabla Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1-B)Z_t$$

Therefore, $\nabla = (1-B)$.

While for a second order difference ;

$$\nabla^2 Z_t = (1-B)^2 Z_t$$

In general, difference of order d will assume the following equation. ;

$$\nabla^d Z_t = (1-B)^d Z_t \tag{10}$$

3.11.4 Stationarity.

Values of a time series at specific time points can be influenced by the underlying trend and seasonal variation, thus making the series non-stationary. For a time series to be stationary the mean, variance and autocovariance should not be a function of time. Non-stationarity in time time series observations can be removed through differencing. If the variance is changing with time, then a log transformation can be applied. Testing for stationarity is done using Augmented Dickey Fuller test (ADF) which tests the null hypothesis of non-stationarity versus the alternative that the observations are stationary.

3.11.5 Autocorrelation function (ACF)

Given a series of observations, $Z_1, Z_2, \ldots, Z_t$ the correlation between two observations say Z_t and an observation at lag k, Z_{t-k}, gives the autoccorrelation function at lag k.

$$\rho_k = corr(Z_t, Z_{t-k}) = \frac{Cov(Z_t, Z_{t-k})}{\sqrt{Var(Z_t)Var(Z_{t-k})}} = \frac{E[(Z_t-\mu_z)(Z_{t-k}-\mu_z)]}{E[(Z_t-\mu_z)^2(Z_{t-k}-\mu_z)^2]}$$

Z_t - time series observations.
Z_{t-k} - is the time series observation at lag k.
μ_Z is the mean.

ACF is thus a collection of ρ_k where $k = 0, 1, \ldots, n$. ACF can be approximated by sample ACF from a time series with finite length. Consider a time series observations of size n; $Z_1, Z_2, Z_3, \ldots, Z_n$. Its sample autocovariance;

$$\Upsilon(k) = \frac{t=1}{n}\sum_{1}^{n}(z_t - \bar{z})(z_{t-k} - \bar{z})$$

Hence, the sample ACF is given by;

$$\widehat{\rho}_k = \frac{\Upsilon(k)}{\Upsilon(0)}$$

Where $K = 0, 1, \ldots, n$

3.11.6 Partial autocorrelation function (PACF)

PACF is a function that gives the correlation between Z_t and Z_{t-1} or Z_t and Z_{t-2} or Z_t and Z_{t-p} while the influence of the intermediate lags is removed.

Time series models works with independent observations which can be assessed by the ACF plots or by use of the Durbin-Watson (DW) test. DW, test the null hypothesis of serial autocorrelation versus the alternative that regression residuals are autocorrelated. For this test α=0.05 was chosen.

The test statistic;

$$DW = \frac{\sum_{k=2}^{n}(\varepsilon_k - \varepsilon_{k-1})^2}{\sum_{k=1}^{n}\varepsilon^2{}_k} \tag{11}$$

where; $\varepsilon_k = Z_k - \widehat{Z}_k$ are the residuals resulting from linear regression while Z_k and $\widehat{Z}_k$ are the observed and predicted values respectively.

3.11.7 Cochrane-Orcutt method.

In the case where positive autocorrelation is remarked from the residuals, then independence can be reached by applying the Cochrane-Orcutt method as described by Cochrane and Orcutt (1949). Iteration procedure is outlined outlined below;

- From the time series observations, model residuals are obtained from a fitted linear regression.
- First order autocorrelation coefficient ρ is computed from the residuals obtained. More precisely as follows;

$$\varepsilon_k = \rho\varepsilon_{k-1} + \omega_k$$

- The original time series values are then transformed by using the autocorrelation coefficient ρ obtained above.

$$Z^* = Z_k - \rho Z_{k-1}$$

- The newly transformed variable is again regressed and the independence of the lagged residuals is assessed.
- The process is repeated until convergence is realized.

3.12 Time series models

3.12.1 Autoregressive (AR) model.

In AR model, a future value is predicted using the linear combination of previous values. The order of AR model gives the number of the lagged values used to estimate the present value. Below is a first-order, AR (1) model.

$$Z_t = \theta_1 Z_{t-1} + \varepsilon_t \tag{12}$$

Where Z_t is the time series values; θ is the regression parameter to be determined; $\varepsilon_t \sim WN(0,\sigma^2)$ and is uncorrelated with Z_s for $s < t$. In equation 12, Z_{t-1} , (the previous response value), is the predictor of the present value, Z_t.

A p^{th} order, AR(p) is a multiple linear regression model having a value of the series at a given time as a linear combination of values at $t-1,\ldots,t-k$. An AR model of order p is given by the following equation;

$$Z_t = \theta Z_{(t-1)} + \ldots + \theta_p Z_{(t-p)} + \varepsilon_t \tag{13}$$

Equation 13 can be simplified as;

$$Z_t = \sum_{k=1}^{p} \theta_k Z_{t-k} + \varepsilon_t \tag{14}$$

Introducing the back-shift operator,

$$Z_t - \sum_{k=1}^{p} \theta_k Z_{(t-k)} = \varepsilon_t$$

$$(1 - \theta_1(B^1) - \theta_2(B^2) -, \ldots, -\theta_p(B^p))Z_t = \varepsilon_t$$

$$\Big(1 - \sum_{k=1}^{p} \theta_k B^k\Big) Z_t = \varepsilon_t$$

$$\theta(B) Z_t = \varepsilon_t \tag{15}$$

3.12.2 Moving Average (MA) model

MA model of order q, is a weighted sum of the present error term and q previous error terms given by the following;

$$Z_t = \varepsilon_t + \rho\varepsilon_{(t-1)} + \rho_2\varepsilon_{(t-2)} + \ldots + \rho_q\varepsilon_{(t-q)} \tag{16}$$

Equation 16 can be simplified as;

$$Z_t = \varepsilon_{\mathrm{t}} + \sum_{k=1}^{q} \rho_{\mathrm{k}} \varepsilon_{\mathrm{t\text{-}k}} \tag{17}$$

Introducing the backshift operator B such that $B\varepsilon_t = \varepsilon_{t-1}$

$$Z_t = (1 + \rho(B) + \rho_2(B^2) +, \ldots + \rho_{\mathrm{q}}(B^{\mathrm{q}}))\varepsilon_{\mathrm{t}}$$

$$Z_t = (1 + \sum_{k=1}^{q} \rho_{\mathrm{k}} B^{\mathrm{k}})\varepsilon_{\mathrm{t}}$$

$$Z_t = \rho(B)\varepsilon_{\mathrm{t}} \tag{18}$$

3.12.3 Autoregressive Moving Average model (ARMA)

Box-Jenkins ARMA is a mixture of the Autoregressive model of order p, AR(p) and a moving average model of order q, MA(q). The resulting model assumes the form ARMA (p, q), where p and q are respectively the orders of AR and MA models. The ARMA (p, q) model is used for analyzng stationary time series. The process Z_t; t $\in \mathbb{Z}$ is an autoregressive moving average of order (p, q) such that $Z_t \sim ARMA(p,q)$ such that;

$$Z_t = \theta_1 Z_{(t-1)} + \ldots + \theta_{\mathrm{p}} Z_{(t-p)} + \varepsilon_{\mathrm{t}} + \rho_1 \varepsilon_{\mathrm{(t\text{-}1)}} + \ldots + \rho_{\mathrm{q}} \varepsilon_{\mathrm{(t\text{-}q)}}$$

$$Z_t = \varepsilon_{\mathrm{t}} + \sum_{k=1}^{p} \theta_{\mathrm{k}} Z_{t-k} + \sum_{j=1}^{q} \rho_{\mathrm{j}} \varepsilon_{\mathrm{t\text{-}j}} \tag{19}$$

where $\varepsilon_{\mathrm{t}} \sim WN(0, \sigma^2)$ and $\theta_1, \theta_2, \ldots, \theta_{\mathrm{p}}, \rho_1, \rho_2, \ldots, \rho_{\mathrm{q}}$ are $(p+q)$ parameters to be determined.
Rewriting equation 19

$$Z_t - \sum_{k=1}^{p} \theta_{\mathrm{k}} Z_{t-k} = \varepsilon_{\mathrm{t}} + \sum_{j=1}^{q} \rho_{\mathrm{j}} \varepsilon_{\mathrm{t\text{-}j}}$$

$$\left[(1 - \theta_1(B^1) - \theta_2(B^2) -, \ldots, -\theta_{\mathrm{p}}(B^{\mathrm{p}})\right] Z_t = \left[1 + \rho(B) + \rho_2(B^2) +, \ldots + \rho_{\mathrm{q}}(B^{\mathrm{q}})\right] \varepsilon_{\mathrm{t}}$$

Hence,

$$\theta(B) Z_t = \rho(B)\varepsilon_{\mathrm{t}} \tag{20}$$

3.12.4 Auto-regressive integrated moving average model (ARIMA)

ARIMA(p, d, q) is an extension of ARMA(p, q) model. Non-stationary ARMA model is made stationary by differencing. The resulting model after differencing is an ARIMA (p, d, q) where d signifies the non-seasonal terms required to stabilize the integrated time series observations. Below is the general form an ARIMA (p, d, q) is

$$\theta(B) \nabla^{\mathrm{d}} Z_t = \rho(B)\varepsilon_{\mathrm{t}} \tag{21}$$

Where

1. $\theta(B) = [1 - \theta(B^1) -, \ldots, -\theta_{\mathrm{p}}(B^{\mathrm{p}})] = \left(1 - \sum_{k=1}^{p} \theta_{\mathrm{k}} B^{\mathrm{k}}\right)$
2. $\nabla^{\mathrm{d}} Z_t = (1-B)^{\mathrm{d}} Z_t$
3. $\rho(B)\varepsilon_{\mathrm{t}} = (1 + \rho(B) + \rho_2(B^2) +, \ldots + \rho_{\mathrm{q}}(B^{\mathrm{q}}))\varepsilon_{\mathrm{t}} = \left(1 + \sum_{k=1}^{q} \rho_{\mathrm{k}} B^{\mathrm{k}}\right)\varepsilon_{\mathrm{t}}$

Hence combining the three equations forms the ARIMA(p,d,q) process.

$$\left(1 - \sum_{k=1}^{p} \theta_{\mathrm{k}} B^{\mathrm{k}}\right)(1-B)^{\mathrm{d}} Z_t = \left(1 + \sum_{k=1}^{q} \rho_{\mathrm{k}} B^{\mathrm{k}}\right)\varepsilon_{\mathrm{t}} \tag{22}$$

3.12.5 Seasonal Autoregressive integrated moving average (SARIMA) model.

Time series observations can have patterns that repeat itself at a regular period of time. Seasonal ARIMA is an extension of non-seasonal ARIMA model for analyzing seasonal non-stationary data. The periodicity in time series can be weekly, quarterly or monthly. The notation of the multiplicative SARIMA is given as ARIMA$(p,d,q)(P,D,Q)_m$; m denotes the period at which the seasonal pattern repeats itself, the parameters p, d and q are non-seasonal integers corresponding to the order of non-seasonal AR, non-seasonal differencing and non-seasonal MA while P, D, and Q are the order of seasonal AR, order of seasonal differencing and the order of seasonal MA respectively.
The mathematical form of SARIMA$(p,d,q)(P,D,Q)_m$ is as follows.

$$\phi_{\mathrm{P}}(B^{\mathrm{m}})\theta(B)\nabla_m^D \nabla^{\mathrm{d}} Z_t = \Psi_{\mathrm{Q}}(B^{\mathrm{m}})\rho(B)\varepsilon_{\mathrm{t}} \tag{23}$$

Z_t is the non-stationary time series observations, ε_t is the white noise and m is the period of the time series.

Notations;

1. $\theta(B) = \left[1 - \theta_1(B^1) - \theta_2(B^2) -, \ldots, -\theta_{\mathrm{p}}(B^{\mathrm{p}})\right]$ is the non-seasonal AR(p)
2. $\phi_{\mathrm{P}}(B^{\mathrm{m}}) = \left[1 - \phi_1(B^{\mathrm{m}}) - \phi_2(B^{2\mathrm{m}}) -, \ldots, -\phi_{\mathrm{P}}(B^{\mathrm{Pm}})\right]$ is the seasonal AR(P)
3. $\rho(B) = \left[1 + \rho_1(B) + \rho_2(B^2) +, \ldots + \rho_{\mathrm{q}}(B^{\mathrm{q}})\right]$ is the non-seasonal MA(q) process.
4. $\Psi_{\mathrm{Q}}(B^{\mathrm{m}}) = \left[1 + \Psi_1(B^{\mathrm{m}}) + \Psi_2(B^{2\mathrm{m}}) +, \ldots + \Psi_{\mathrm{Q}}(B^{\mathrm{Qm}})\right]$ is the seasonal MA(Q)
5. $\nabla^{\mathrm{d}} = (1-B)^{\mathrm{d}}$ denote the non-seasonal difference
6. $\nabla^{\mathrm{D}} = (1-B^m)^{\mathrm{D}}$ denote the seasonal difference

3.13 Box-Jenkins model building steps.

Box-Jenkins (1970) describes a procedure towards modelling ARIMA models. The steps involved are discussed in the next subsections.

3.13.1 Model identification.

Box-Jenkins model works with the notion that the data is a realization of a stationary process; hence the first step is to determine the stationarity and also the possibility of seasonality in the time series. Time series observations that fluctuate with a constant variance around a constant mean can be supposed to be stationary. The stationarity of the process generating the data can be examined by plotting the time series values. Seasonality in the data can be detected by an auto-correlation plot. Transformation of the time series by differencing is necessitated if it is found to defy the assumption of stationarity.

Once the stationarity has been achieved, the order (p) of AR model, differencing order d, MA order (q) are identified with the corresponding seasonal values of P, D, and Q. ACF plots and the PACF plots are used in identifying the order of these parameters. The order of the non-seasonal AR (p) and MA (q) are chosen according to the characteristics shown in table 1. Similarly, the values of seasonal AR and MA processes are obtained from the

Table 1. Order of AR and MA processes

Model	ACF	PACF
1. AR(p)	The lags decay exponentially	cuts off after p lags
2. MA(q)	Cuts off after q lags	The lags decay exponentially
3. ARMA(p, q)	Tails off	Tails off

ACF and the PACF plots. For instance an ARIMA $(0,0,0)(1,0,0)_{12}$ represents a seasonal part of an AR process of order $(P = 1)$ which is having a significant PACF spike at lag 12 and an exponentially decaying seasonal lags of the auto correlation function. While an $ARIMA(0,0,0)(0,0,1)_{12}$ represents a seasonal part of the MA process of order (Q = 1) having a significant ACF spike at lag 12 and an exponentially decaying seasonal lags of the partial auto correlation function. Presence of a seasonal component in the data will require seasonal differencing to enhance stationarity of the time series.

Model identification will always lead to the selection of more than one probable models. The most parsimonious model is arrived at by checking the adequacy of the models. The adequacy can be assessed by comparing the Akaike information Criteria (AIC) or the Schwarz Bayesian Information Criteria (BIC). The model that presents minimum values

of AIC and BIC is considered the most adequate. The statistical formulae for the two criteria approaches are given by;

$$\text{AIC} = -2log(L) + 2v$$

and

$$\text{BIC} = -2ln(L) + log(m)v$$

L represents the maximum likelihood, and v the number of model parameters.

3.13.2 Parameter estimation.

Estimation of the model parameters is then done once a tentative model has been identified. The estimation process can be done in several ways; the method of least squares can be applied, the approach of the maximum likelihood estimation (MLE) can be utilized or even the Yule-Walker method. For the purpose of this study we discuss the MLE approach.

Maximum likelihood estimation (MLE).

MLE method is by far the most widely used approach for estimating the unknown parameters of a probability distribution. The principle of this approach considers maximizing the likelihood function so that the estimated parameters are values for which the observed data is most probable.

Consider an AR(1) process which is given by;

$$Z_t = \psi + \theta Z_{t-1} + \varepsilon_t$$

Where $\varepsilon_t \sim i.i.d\ N(0, \sigma^2)$, under the MLE approach, the parameters to be estimated include $\phi = (\psi, \theta, \sigma^2)$.

Consider Z_1 as the first observation. Its mean and variance can be shown to be;

$$E(Z_1) = \mu = \psi/(1-\theta) \text{ and } E(Z_1 - \mu)^2 = \sigma^2/(1-\theta^2)$$

The probability density distribution for Z_1 is given by:

$$f_{Z_1}(Z_1;\phi) = \left(2\pi\sigma^2/(1-\theta^2)\right)^{-\frac{1}{2}} exp - \frac{1}{2}\left[\frac{\left(Z_1 - \psi/(1-\theta)\right)^2}{\sigma^2/(1-\theta^2)}\right] \tag{24}$$

Also consider a second a second observation Z_2 which is conditional on $Z_1 = z_1$ such that

$$Z_2 = \psi + \theta Z_1 + \varepsilon_2$$

This implies that the first two terms of the equations i.e ψ and θZ_1 are constants. The error term $\varepsilon_2 \sim (0, \sigma^2)$. Thus $(Z_2|Z_1 = z_1) \sim N(\psi + \theta Z_1, \sigma_2)$.

$$f_{Z_1|Z_2}(Z_2|Z_1;\phi) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} exp - \frac{1}{2}\left[\frac{\left(Z_2 - \psi - \theta Z_1)\right)^2}{\sigma^2}\right] \tag{25}$$

Thus, the joint density of Z_1 and Z_2 is the product of equation 24 and equation 25.

$$f_{Z_2,Z_1}(Z_2,Z_1;\phi) = f_{Z_2|Z_1}(Z_2|Z_1;\phi) \cdot f_{Z_1}(Z_1;\phi)$$

Following the same pattern, the distribution of Z_3 conditional on Z_1 and Z_2 is given by;

$$f_{Z_3|Z_2,Z_1}(Z_3|Z_2,Z_1;\ \phi) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} exp - \frac{1}{2}\left[\frac{\left(Z_3 - \psi - \theta Z_2)\right)^2}{\sigma^2}\right] \tag{26}$$

and the joint density of Z_1, Z_2 and Z_3 is the product of their conditional distributions.

$$f_{Z_3,Z_2,Z_1}(Z_3,Z_2,Z_1;\ \phi) = f_{Z_3|Z_2,Z_1}(Z_3|Z_2,Z_1;\ \phi) \cdot f_{Z_2,Z_1}(Z_2,Z_1;\ \phi) \tag{27}$$

It follows that the density of the last observation Z_t conditional on its previous observations $Z_1, Z_2, \cdots, Z_{t-1}$ is written as;

$$f_{Z_t|Z_{t-1},Z_{t-2}\cdots,Z_1}(Z_t|Z_{t-1},Z_{t-2},\cdots,Z_1;\phi) =$$

$$\left(2\pi\sigma^2\right)^{-\frac{1}{2}} exp - \frac{1}{2}\left[\frac{\left(Z_t - \psi - \theta Z_{t-1})\right)^2}{\sigma^2}\right] \tag{28}$$

For t distributions, the density is given by

$$\begin{aligned} &f_{Z_t,Z_{t-1},Z_{t-2},\ \cdots,\ Z_1}(Z_t,Z_{t-1},Z_{t-2},\cdots,Z_1;\phi) = \\ &f_{Z_t|Z_{t-1}}(Z_t|Z_{t-1}\ ;\phi) \cdot f_{Z_{t-1},Z_{t-2}\cdots,Z_1}(Z_{t-1},Z_{t-2},\cdots,Z_1\ ;\phi) \end{aligned} \tag{29}$$

The likelihood function

$$\begin{aligned} &f_{Z_t,Z_{t-1},Z_{t-2},\cdots,Z_1}(Z_t,Z_{t-1},Z_{t-2},\cdots,Z_1;\ \phi) = \\ &f_{Z_1}(Z_1;\ \phi) \cdot \prod_{t=2}^{t=n} f_{Z_t,Z_{t-1}}(Z_t|Z_{t-1};\ \phi) \end{aligned} \tag{30}$$

The log-likelihood function $L(\phi)$ is thus;

$$L(\phi) = \log f_{Z_1}(Z_1;\phi) + \sum_{t=2}^{t=n} \log f_{Z_t|Z_{t-1}}(Z_t|Z_{t-1};\ \phi) \tag{31}$$

Substituting equation 31 with 24 and 28 gives the log-likelihood of AR(1) process

$$L(\phi) = \log\left\{\left(2\pi\sigma^2/(1-\theta^2)\right)^{-\frac{1}{2}} exp - \frac{1}{2}\left[\frac{\left(Z_1 - \psi/(1-\theta)\right)^2}{\sigma^2/(1-\theta^2)}\right]\right\} + \sum_{t=2}^{t=n} \log\left\{\left(2\pi\sigma^2\right)^{-\frac{1}{2}} exp - \frac{1}{2}\left[\frac{\left(Z_t - \psi - \theta Z_{t-1})\right)^2}{\sigma^2}\right]\right\} \tag{32}$$

$$L(\psi,\theta,\sigma^2) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\left[\sigma^2/(1-\theta^2)\right] - \frac{1}{2}\left[\frac{\left(Z_1 - \psi/(1-\theta)\right)^2}{\sigma^2/(1-\theta^2)}\right] - \frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2) - \frac{1}{2}\sum_{t=2}^{t=n}\left[\frac{(Z_t - \phi - \theta Z_{t-1})^2}{\sigma^2}\right] \tag{33}$$

3.13.3 Diagnostic check of the model

Model evaluation is done by analyzing the residuals of the resulting fitted model. An adequate model has its residuals follow a white noise process. There are several approaches that can be used to evaluate the adequacy of the model. Plots of autocorrelation and Partial autocorrelation functions can be used to asses the independence of residual series. For residuals to exhibit independence of observations, the autocorrelation coefficients for lags greater than 1 should lie within a defined threshold. For simplicity,the coefficients should not be significantly different from zero, contrary to this signifies presence of a systematic pattern in the residuals.

The Ljung-Box test is yet another approach to asses the lack of fit of the model. The statistic test the null hypothesis that the residuals follow a white noise process. ;

$$Q = n(n+2)\sum_{m=1}^{k}(t-m)^{-1}\,\hat{r}_m^2$$

The statistic $Q \sim \chi^2_{1-\alpha,\,k-p-q}$; p and q are the corresponding orders of AR and MA processes, t is the size of the time series, $\hat{r}_m$ gives the estimated autocorrelation coefficient at lag m and k represents the number of autocorrelations tested. When $Q > \chi^2_{1-\alpha,\,k-p-q}$ then the model being tested exhibits a lack of fit.

3.13.4 Forecasting

Once a parsimonious model has been identified, the final stage under Box-Jenkins involves the prediction of future values. Assuming an ARMA(p, q) given below;

$$\theta(B)Z_t = \rho(B)\varepsilon_{\mathrm{t}}$$

$$Z_t = \theta_1 Z_{(t-1)} + \ldots + \theta_p Z_{t-p} + \varepsilon_t + \rho_1 \varepsilon_{(t-1)} + \ldots + \rho_q \varepsilon_{(t-q)}$$

Taking the present time as t hence the h period forecast is given by;

$$Z_{t+h} = \theta_1 Z_{(t+h-1)} + \ldots + \theta_p Z_{(t+h-p)} + \varepsilon_{(t+h)} + \rho \varepsilon_{(t+h-1)} + \ldots + \rho_q \varepsilon_{(t+h-q)}$$

The parameter values on the right of the equation are estimated and replaced.

3.14 Analysis of Seasonal variation.

As defined earlier, seasonality is the periodic pattern that reoccurs at intervals of time within the year. To evaluate the seasonal variations in LBW prevalence, we used the Moving Average (MA) method to obtain the seasonal indices associated with the months of the year. Below is the description of the steps used to obtain the indices.

- First, we obtained a 12-monthly moving average of LBW prevalence, then followed by a two-monthly moving average to center the data. Herein, the aim is to smoothen the trend component, thus removing seasonality and the random component.

  The 2×12-MA can be illustrated as follows;

$$T_t = \frac{1}{2}\left[\frac{1}{12}\left(Z_{(t-6)} + Z_{(t-5)} + \ldots + Z_t + Z_{(t+1)} + Z_{(t+2)} + \ldots + Z_{(t+5)}\right) + \frac{1}{12}\left(Z_{(t-5)} + \ldots + Z_t + Z_{(t+1)} + \cdots + Z_{(t+6)}\right)\right]$$

  Simplifying the above equation yielded the following;

$$T_t = \frac{1}{24}Z_{(t-6)} + \frac{1}{12}\left(Z_{(t-5)} + \ldots + Z_t + Z_{(t+1)} + \ldots + Z_{(t+5)}\right) + \frac{1}{24}Z_{(t+6)} \qquad \textbf{(34)}$$

- To obtain the seasonal variations, the smoothed trend values are subtracted from the original series as follows;

$$Z_t - T_t = S_t + \varepsilon_t$$

  where Z_t is the original prevalence values, T_t is the smoothened trend values while S_t and ε_t are the seasonal and noise respectively (detrended series).

- The residual component is then removed by averaging the detrended values on monthly basis hence obtaining the unadjusted seasonal variations.

- Finally, the adjusted Seasonal variations were obtained by subtracting the average of the unadjusted seasonal variations (adjustment factor) from each of the unadjusted monthly seasonal variations. The adjustment herewith is to make seasonal indices sum up to zero.

3.15 Malaria endemicity.

In evaluating the level of malaria transmission in a population Brabin et al. (1999) developed a birthweight nomogram, Figure 2. It is a simple tool constructed to monitor transmission of malaria and to study the efficacy of interventions aimed at controlling malaria during pregnancy in Afica.

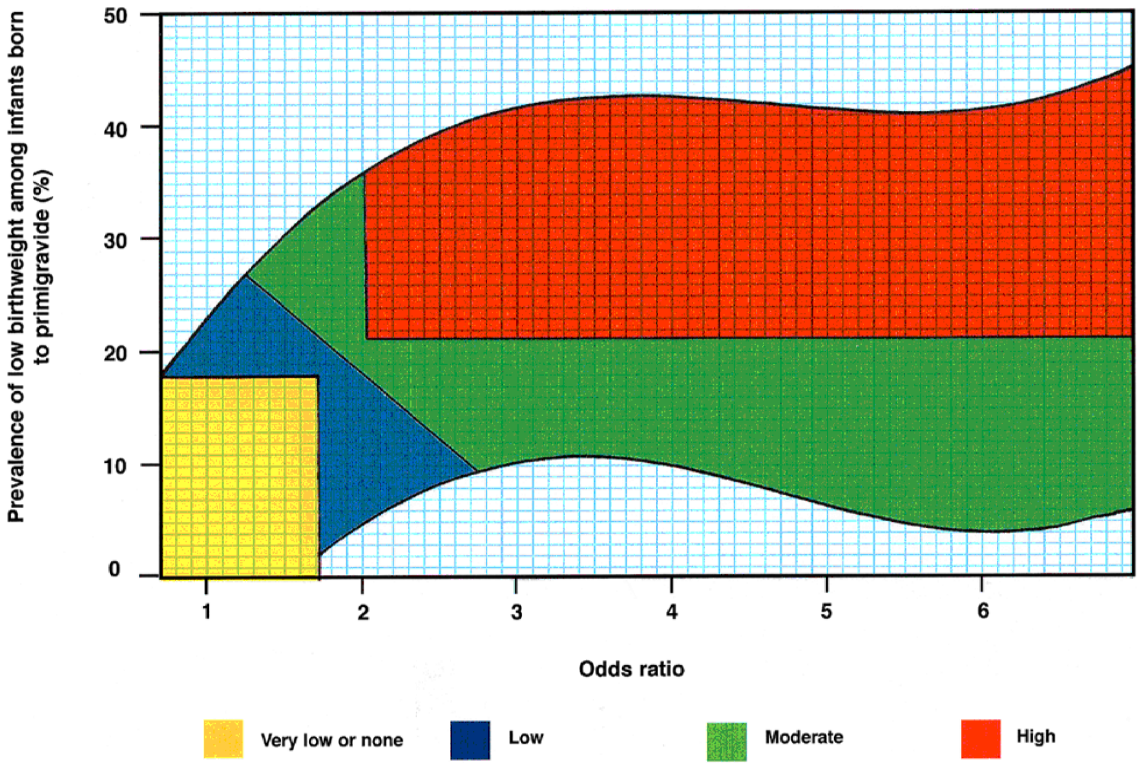

Figure 2. birthweight nomogram indicating transmission intensities.
Source: Brabin et al., 1999

The birth-weight nomogram was designed to show the link between the odds ratio of excess LBW risk and the LBW prevalence among the primigravdae. According to (Brabin et al., 1999) the odds ratio for excess low birth weight risk signified the excess risk attributed to malaria whereas low birthweight prevalence was attributed to the excess risk from all other causes. An OR of 1.7 signified a cutting-off point between malaria and non-malarious areas because of its 90 percent sensitivity and 90 percent specificity for detecting a population exposed to malaria, whilst an OR lower than 1.7 suggested lower risk of malaria.

In that respect we used the odds ratios of LBW among the primiparity with reference to multiparity to study the epidemiological transition of malaria in kenya. Additive logistic regression model with two independent variables (parity and season) and a binary

dependent variable (birth weight) was used to obtain the odds ratios.

$$\text{Birthweight} = \begin{cases} 1, & \text{Low birth weight} \\ 0, & \text{Normal weight} \end{cases}$$

$$\text{Parity} = \begin{cases} 1, & \text{Primiparity} \\ 2, & \text{Multiparity} \end{cases}$$

$$\text{Season} = \begin{cases} 1, & \text{Wet} \\ 2, & \text{Dry} \end{cases}$$

The general form of the equation was defined as follows

$$ln(\frac{\theta}{1-\theta}) = b_0 + b_1 Z_1 + b_2 Z_2 \tag{35}$$

Where;
θ is the probability of low birthweight.
$1-\theta$ is the probability of normal weight.
$b_{j's}$ are the regression coefficients to be estimated; for j=0,1,2.
Z_1 and Z_2 represents parity and season respectively.

Hence, from the equation 35 , we can obtain the odds of low birthweight as

$$\frac{\theta}{1-\theta} = exp(b_0 + b_1 Z_1 + b_2 Z_2) \tag{36}$$

And the corresponding probability of low birth weight;

$$\theta = exp(b_0 + b_1 Z_1 + b_2 Z_2) - \theta[exp(b_0 + b_1 Z_1 + b_2 Z_2)]$$

$$\theta[1 + exp(b_0 + b_1 Z_1 + b_2 Z_2)] = exp(b_0 + b_1 Z_1 + b_2 Z_2)$$

hence,

$$\theta = \frac{exp(b_0 + b_1 Z_1 + b_2 Z_2)}{1 + exp(b_0 + b_1 Z_1 + b_2 Z_2)} \tag{37}$$

Odds ratio is given by $exp(b_j)$.

3.16 Spatial Pattern analysis.

Spatial patterns exhibited in the distribution of certain phenomena are determined by the arrangement of the individual entities in space and the geographical relationships among them. The pattern portrays the underlying spatial process at a particular period of time.

3.16.1 Cluster analysis.

A spatial cluster in a certain location refers to grouping of events or spatial objects. Spatial cluster analysis encompasses methods and algorithms used to group spatial objects in a way that there is maximal resemblance between objects belonging to the same group and minimal otherwise. The analysis helps to unravel the underlying spatial distribution of occurrences and possible correlation with the environmental factors. Spatial cluster modelling approaches have seen considerable application in spatial epidemiology, crime analysis and disease surveillance.

3.16.2 Spatial autocorrelation.

Spatial autocorrelation is simply the correlation between the values of a variable in one location and values of the same variable in the neighboring locations. Positive spatial autocorrelation implies similarity in nearby values such that high values tend to cluster together and low values being near low values.

3.16.3 Spatial pattern detection.

Clustering of events can be detected by quantifying the degree in which the spatial objects cluster themselves in geographical location. Such testing is made possible by setting a hypothesis and applying available methods. Detection involves assessing either global or local clustering of cases or controls.

3.16.4 Spatial scan Statistic.

The spatial scan statistic (SaTScan) Kulldorff (1997) is an open source software and has found a great application in the field of epidemiology. Its usage being mainly to detect and identify the locations of hotspots. Unlike the global techniques which assume homogeneity of the study location, SaTScan adjust for population heterogeneity of different locations. For a specified geographical location, the scan statistic operates by implementing multiple scan windows or circles of varying sizes. Each imposed circle is a potential hotspot whose significance must be tested using likelihood ratio approach. The size of the circle has a default setting of 50% of the population size. For each circle, cases of LBW are counted and the expected number is obtained. Specifying Bernoulli probabil-

ity model, the maximum likelihood statistic for each circle is obtained by comparing the prevalence of LBW in the circle to the outside prevalence. The window with the maximum value becomes the most likely hotspot. The significance of the clusters are then tested for spatial randomness. The test is implemented by Monte Carlo simulations with replications specified by the user, (a default value is 999 replications). In each simulation a likelihood ratio statistic is computed. Then a p value to evaluate the randomness of cases in the cluster is obtained by comparing actual value from likelihood ratio with simulated values. A cluster is inferred if the associated $p < 0.05$.

We implemented our analysis using the Bernoulli model which requires the definition of cases and controls. We set the low birthweights as cases and normal weights was controls. The maximum size of the proportion in each window was set to 50% of the population and 999 Monte Carlo simulations were set to test the significance of each candidate cluster.

Bernoulli principle.

$$\left(\frac{z}{m}\right)^{z}\left(\frac{m-z}{n}\right)^{m-z}\left(\frac{Z-z}{M-m}\right)^{Z-z}\left(\frac{(M-m)-(Z-z)}{M-m}\right)^{(M-m)-(Z-z)} I() \qquad \textbf{(38)}$$

Where;

- z: number LBW cases in a scanned window,
- m: number of LBW cases and controls(Normal weights) in a scanned window,
- Z: Total number of LBW cases
- M: Total number of LBW cases and controls.

4 DATA ANALYSIS AND RESULTS

4.1 Introduction

The sections presents utilizes the methodologies discussed in chapter three. Descriptive statistics forms the first part of the chapter. Later on inferential statistics are presented. Assumptions for the methods used were checked and analysis done. The results are documented in form of tables, figures, bar charts and maps.

4.2 Demographic characteristics

Details of information obtained and included in this study is presented in Figure 3 . Out of 45,749 entries of live singleton births; 8,381 (18.32 %) were excluded because they lived outside KHDSS. Among those who lived within KHDSS, 1,933 (5.17 %) had missing information on weight and were excluded. The remaining 35,435 records of live births were used for analysis and it comprised of 16,964 (48.24 %) females and 18,205 (51.76 %) males.

Maternal information available on the records was also extracted concurrently with that of the child. Out of 35,435 mothers, 59.94% were multiparous, 38.55% were primiparous while 1.53% had their parity status not indicated. Their age distribution was as follows: 17.2% were less than 20 years, 30.76% were 20-24 years, 25.77% were 25-29 years, 15.74% were 30-34 years while 10.53% were above 34 years. The age distribution of multiparous versus primiparous group was as shown in Figure 4.

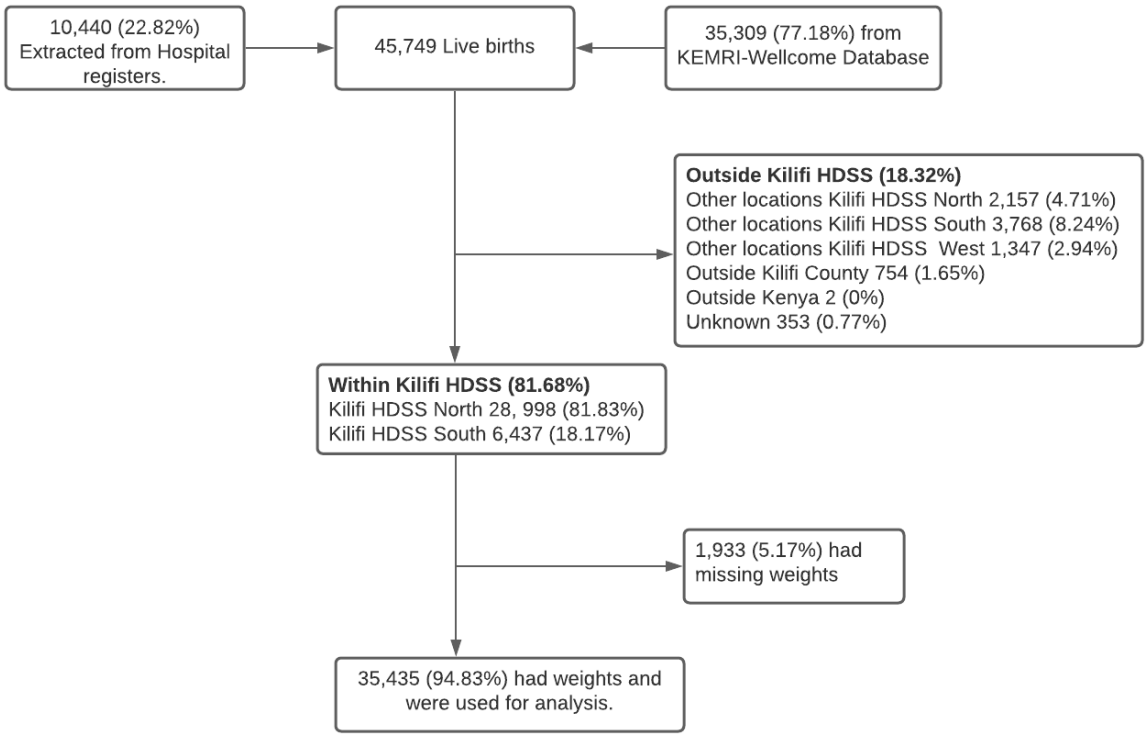

Figure 3. The flow of inclusion criteria of infants in the study.

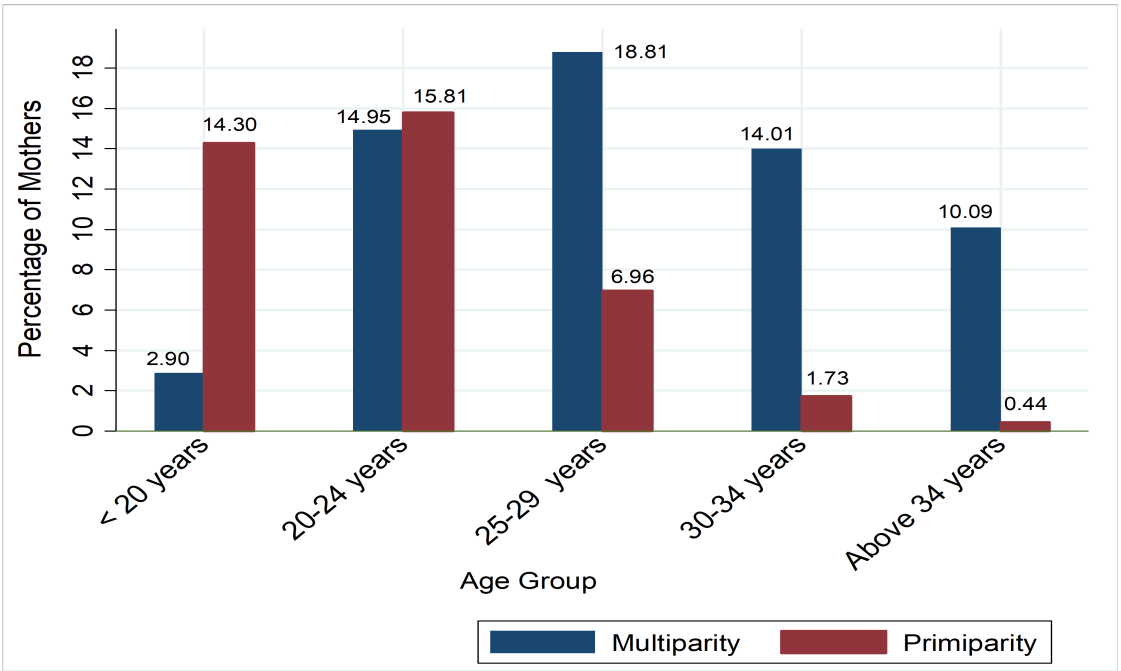

Figure 4. Age distribution of multiparity versus primiparity.

4.3 Birthweight.

The overall mean weight was 2.98 Kg and was higher among males (3.03 Kg) compared to females (2.93 Kg), Table 2. The mean difference in the weights was statistically significant (t statistic = -18.479, p-value = 0.0001). Birthweight by parity also showed a significant difference (t statistic= 23.817, p-value = 0.0001); which was higher among multigravid (3.03 Kg) compared to primigravid (2.90 Kg). Mean weight in Kilifi DSS North (2.989 Kg) and South (2.927 Kg) was relatively same, Table 2 .

4.4 Low birthweight.

The overall prevalence of low birthweight throughout the study period (2006-2019) was 14.23 % and was significantly different among the infants born to primigravids (17.34 %) compared to those of multigravida (12.23 %), p value=0.0001, (Table 2). In the two regions, the prevalence was 17.54% in Kilifi DSS South and 13.50% in Kilifi DSS North. The difference in the two proportions was significant, p value = 0.0001, Table 2

Table 2. Mean weights, LBW prevalence and difference in LBW proportions

	Number (%)	Mean Weight	LBW Prevalence (%)	Difference in LBW proportion, p-value
Gender				
Female	16,964 (48.24)	2.925	15.93	0.0001
Male	18,205 (51.76)	3.028	12.63	
PARITY				
Primiparity	21,239 (60.87)	2.895	17.34	0.0001
Multiparity	13,654 (39.13)	3.031	12.23	
REGION				
Kilifi DSS South	6,437 (18.17)	2.989	17.54	0.0001
Kilifi DSS North	28,998 (81.83)	2.985	13.50	

Focusing on the locations within the two regions, Table 3 , LBW prevalence ranged from 12.50% in Kilifi Township an area in the Northern region to 21.58% in Banda ra Salam an area in the Southern region. The corresponding mean weights of the two locations were 3.03 Kg and 2.87 Kg respectively. The low prevalence of LBW in Kilifi township could be explained by the close proximity to the health care facilities within the urban area.

Table 3. LBW prevalence at locational level.

Location	Mean Birthweight (Kg)	LBW prevalence (%)
Kilifi DSS North		
Gede	2.973	13.51
Kilifi Township	3.028	12.50
Matsangoni	2.885	19.11
Ngerenya	2.957	13.99
Roka	2.926	15.77
Sokoke	2.934	15.05
Takaungu / Mavueni	2.946	14.75
Tezo	2.974	13.27
Kilifi DSS South		
Banda ra Salam	2.865	21.58
Chasimba	2.936	16.14
Jaribuni	2.988	12.95
Junju	2.926	18.54
Kauma	2.935	16.37
Mtwapa	2.927	17.86
Ziani	2.926	16.04

4.4.1 Low birthweight trends.

Prevalence of LBW on a four year period, Table 4, seemed to have a downward trajectory from 2014. Between 2006 and 2009 the prevalence was 14.24%, between 2010 and 2013 it rose to 16.56%.A drop was observed with a record of 13.80% during the period of 2014 and 2017 while between 2018 and 2019 (2-year period) it further decreased to 12.75%.

Table 4. prevalence of low birthweight on a 4-year basis.

Year category	Live Births (%)	Mean weight	Low Birthweight (%)
2006-2009	6,220 (17.55)	2.987	14.24
2010-2013	7,863 (22.19)	2.925	16.56
2014-2017	12,246 (34.56)	2.985	13.80
2018-2019	9,106 (25.70)	3.01	12.75
2006-2019	**35,435 (100)**	**2.978**	**14.23**

Figure 5 shows the pattern of LBW prevalence for the period of 2006 - 2019 in Kilifi HDSS - (North and South), Kilifi HDSS North and Kilifi HDSS South.

The overall prevalence (both parities combined) in the whole region showed a decreasing rate in the first three years which later rose to around 2012. A downward movement was then observed from 2013. The highest prevalence was recorded in 2012 (17.06%) and the lowest rate in 2017 (11.96%). Low birthweight among the primiparity group maintained a higher prevalence rate throughout the years as compared to the multiparity. In both groups a declining trend was observed though the manifestation occurred in different years, with multiparity exhibiting a decline from 2011 while the primiparity after 2012. The LBW ratios of primiparity in relation to multiparity seemingly had the same trajectory to that assumed by LBW prevalence among the primiparity.

In Kilifi HDSS North a declining trend in the prevalence among the multiparity was observed begining 2011 while that of primiparity indicated arising trend from 2009 upon which a decline in the prevalence was observed after 2012.

In the southern region a downward trajectory among the primiparity was observed after 2010 while among the multiparity the decline begun after 2009. In both regions LBW prevalence among primiparity was higher compared to that of multiparity.

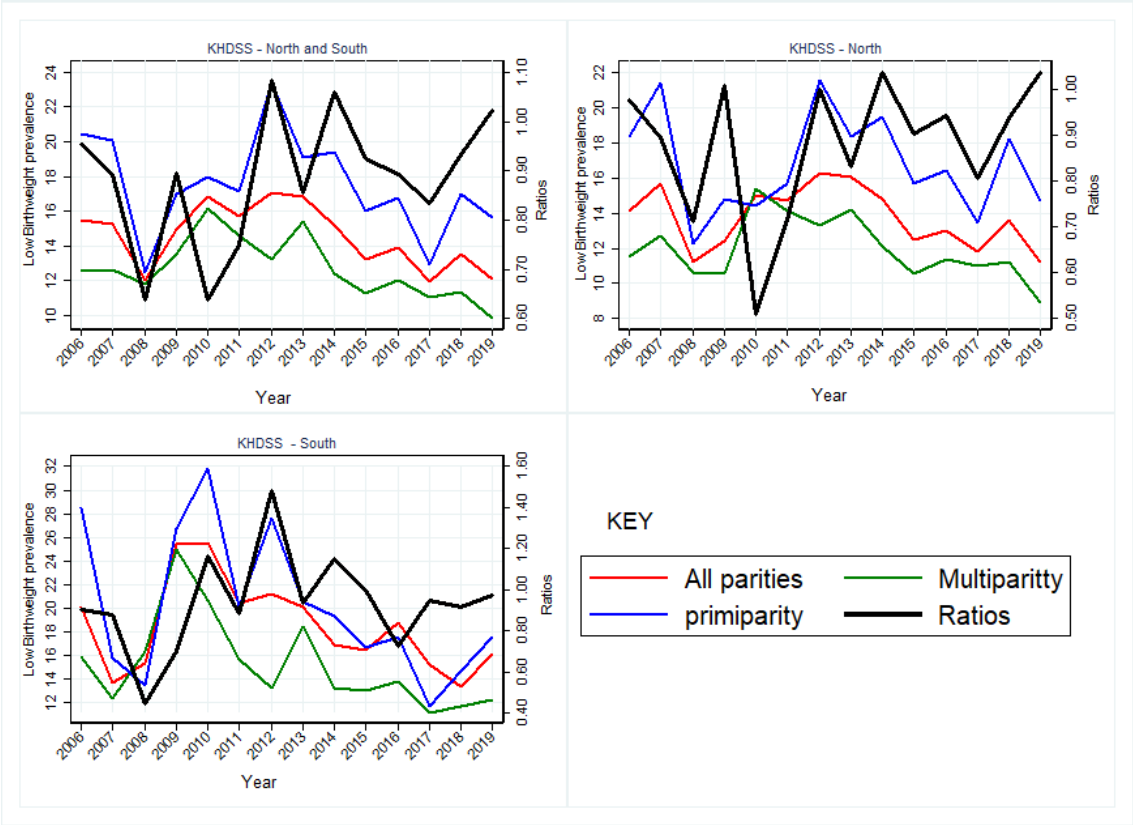

Figure 5. Trends of low birthweight prevalence in all parities combined and differently among the primiparity and multiparity. LBW ratios of primiparity to that of multiparity is displayed on a different scale.

In the two regions (North and South of Kilifi HDSS), LBW prevalence was significantly different, p-value = 0.0001, Table 2 . The difference could clearly be seen in figure 6 . An eyeballing indicated a higher prevalence almost throughout the year in Kilifi HDSS South. The highest prevalence rates were 16.25% (2012) and 25.47% (2009) in North and South respectively while the lowest prevalence were 11.19% (2019) and 13.42% (2018) in North and South respectively.

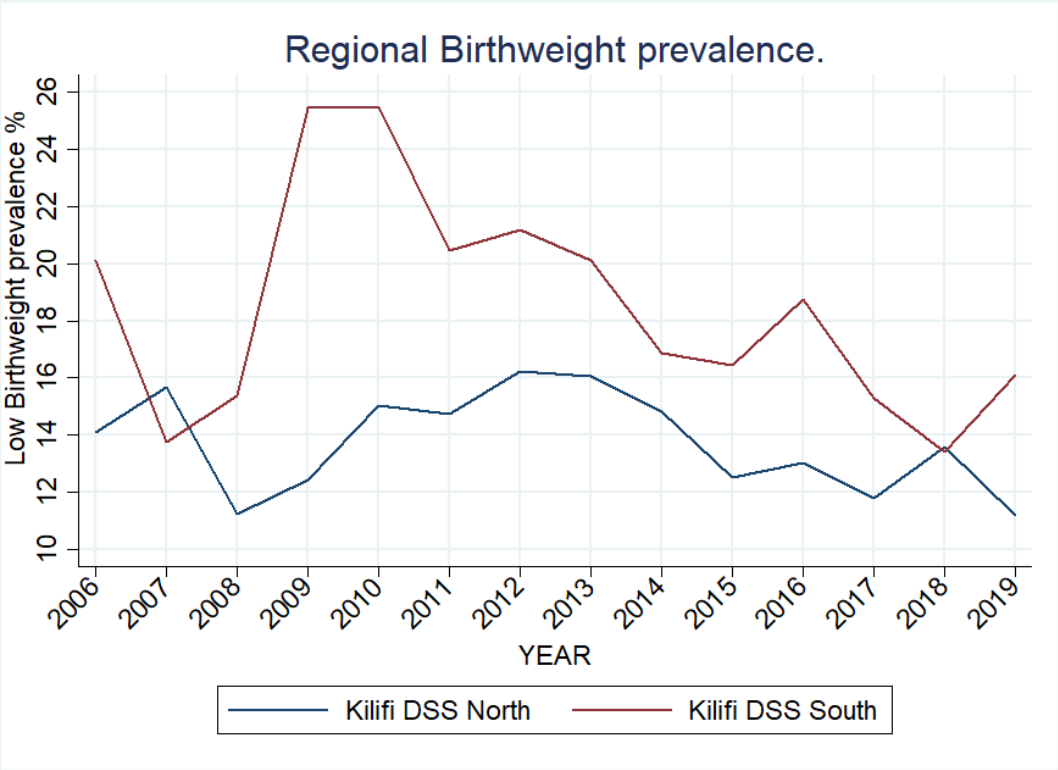

Figure 6. Yearly trend in low birthweight prevalence per Region

Monthly LBW prevalence exhibited a periodic up and down pattern across the months. The month of October had the lowest prevalence of LBW (12.24 %) while April had the highest prevalence (15.82 %) (Table 5). A decreasing trend in LBW prevalence was also observed over the months of the year.

The odds of delivering low a birthweight infant among the primiparous group remained significant throughout the months as compared to multiparous group except in the month of February, where primiparity group in comparison to multiparity group were 19 % more likely to deliver a LBW infant, but that was insignificant. (Table 5).

Table 5. Aggregated monthly low birthweight prevalence and OR of LBW among primiparity compared to multiparity.

Month	Number(n)	primiparity (%)	multiparity (%)	LBW (n, %)	OR (95 % CI)
January	2,436	39.72	60.28	367 (15.36)	1.44 (1.15 1.81)
February	2,569	36.44	63.56	375 (14.48)	1.19 (0.95 1.49)
March	3,654	37.41	62.59	506 (14.26)	1.53 (1.27 1.86)1
April	3,868	39.29	60.71	594 (15.82)	1.38 (1.16 1.65)
May	3,682	40.32	59.68	511 (14.31)	1.49 (1.24 1.80)
June	3,057	38.33	61.67	454 (15.24)	1.80 (1.47 2.21)
July	3,025	38.34	61.66	444 (14.84)	1.54 (1.25 1.89)
August	2,672	39.40	60.60	356 (13.47)	1.52 (1.22 1.91)
September	2,618	41.57	58.43	379 (14.56)	1.57 (1.26 1.95)
October	2,850	41.34	58.66	351 (12.24)	1.50 (1.20 1.88)
November	2,821	39.66	60.34	385 (14.27)	1.58 (1.27 1.97)
December	2,183	37.62	62.38	321 (14.88)	1.65 (1.30 2.10)

4.5 Trend significance and change point analysis of LBW prevalence.

In this section we evaluate the significance of the long-term trend in the prevalence of LBW data using MK test. Results on trend significance can be seen in section 4.5.1 below.

First, we checked on the prerequisites of the test. MK test result is significantly affected by the presence of periodicities and serial autocorrelation in a time series data. To check the two aspects, we plotted the ACF to check for the autocorrelation and also a plot of the decomposed time series to assess the seasonal component. From the two plots, it was clear that the data was serially autocorrelated due the presence of significant spikes at different lags, Figure 7 . The decomposed series in figure 8 exhibits a regular seasonal pattern.

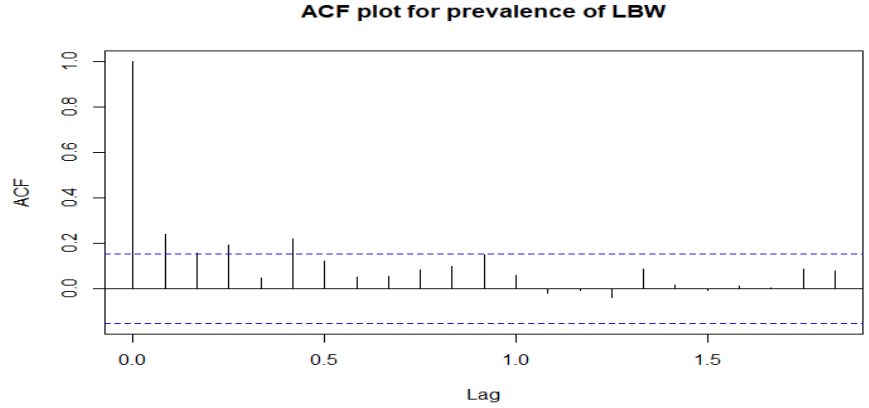

Figure 7. Auto-correlation plot for prevalence of LBW

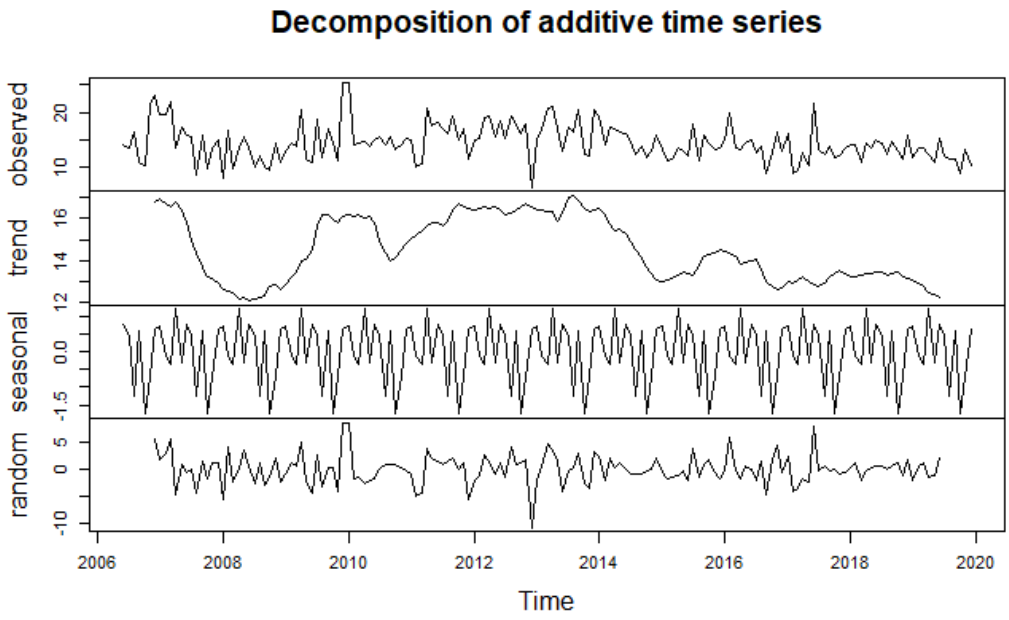

Figure 8. Classical decomposition of the low birthweight series.

We de-seasonalize the LBW series by subtracting the seasonal component from the original series. i.e $Z_t - S_t = T_t + \varepsilon_t$, where Z_t is the original series, S_t is the seasonal component, T_t is the trend and ε_t is the noise. To complement the ACF outcome of serial auto-correlation, we further coupled it with DW test to assess the independence of observations. DW, tests the null hypothesis of no serial auto-correlation against the alternative that the regression residuals are autocorrelated. The results indicated a positive autocorrelation; $DW = 1.545, P\ value = 0.002$.
To remove the positive autocorrelation, we used the Cochrane Orcutt method as explained in section 3.11.7. We obtained an estimated correlation coefficient ρ by assessing the independence of residual terms of fitted linear regression model. From the DW test, the estimated value was, $\rho = 0.2217$. The obtained ρ value was then used to transform the original variable values; $Z_t^{'} = Z_t - 0.2217\ Z_{t-1}$. Again, we tested the independence of the newly transformed variable by assessing the residuals obtained by running the regression once again. Results from the DW indicated absence of autocorrelation, $DW = 2.022, P\ value = 0.984$..

4.5.1 Mann-Kendall trend outcome.

The rank-based MK test indicated a significantly decreasing trend in the LBW prevalence over time, $Z_s = -0.116, P\ value < 0.05$. Mann Kendall test in itself detects a trend and gives the direction but does not quantify its magnitude. To estimate the trend's magnitude, we used the Sen's slope approach. The result from this approach indicated a significant slope, / $S = -0.0118, P\ value < 0.05$.

The decreasing trend from MK test can be visualized from the fitted locally weighted least squares (lowess) regression on the time series observations, Figure 9.

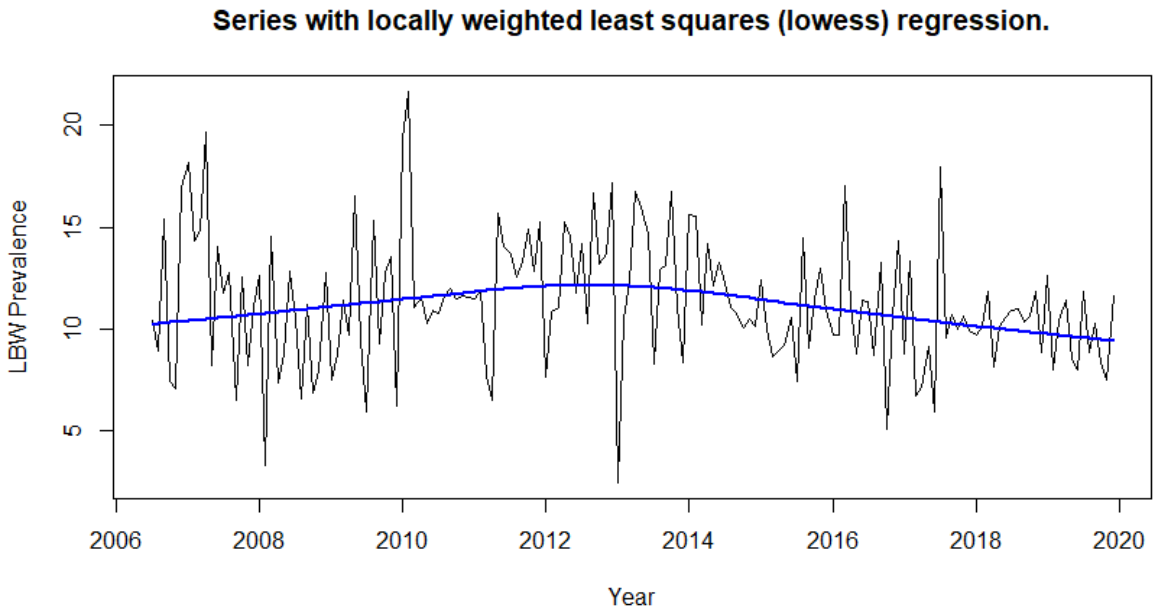

Figure 9. General trend in low birthweight prevalence over time.

4.5.2 Change point analysis (CPA).

CPA involves the estimation of points and locations of a time series where statistical properties change abruptly. In our analysis, we let **t** to be the location at which the mean prevalence of LBW begin to change. To detect this change, we tested the null hypothesis for unchanging mean prevalence in both parities combined against the alternative hypothesis of presence of a change-point in the LBW prevalence at location **t**.

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n$
$H_1 : \mu_1 = \cdots = \mu_t \neq \mu_{t+1} = \cdots = \mu_n$

While testing the two hypotheses, we used the likelihood ratio (LR) approach to compare models presented by the two hypotheses as explained in section 3.10. From this analysis a significant changepoint at around 2014 was detected, (Figure 10). This signified an abrupt change in the prevalence of LBW after the period 2006 to mid-2014. The estimated LBW prevalence in the first segment was 11.92 % after which it dropped to 10.32 % in the remaining time period (between mid-2014 and 2019).

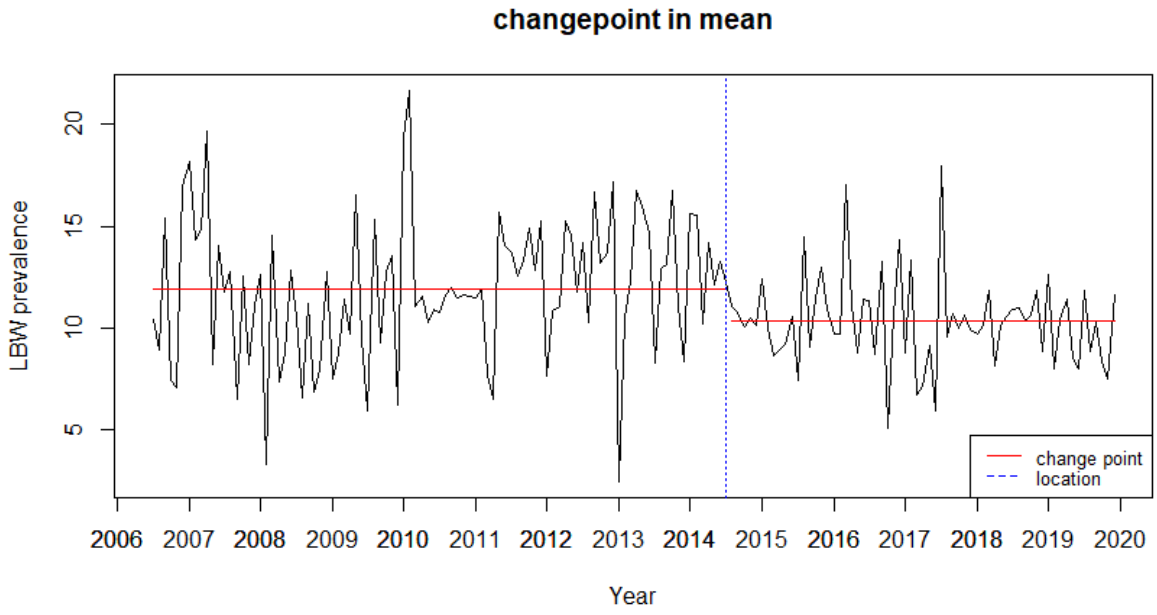

Figure 10. Changepoint in the prevalence of Low Birthweight.

4.6 Interaction between trend and change point.

The obtained changepoint divide the time series into (q+1) segments. MK test has been used to check for possible interaction of trend with the change point. Considering the detected change-point, Figure 10, we divided the series into two segments and tested for trend in each segment. MK test could detect a significant increasing trend in the first segment, Z_s= 0.157, p value = 0.0240. The second segment had a decreasing trend but was not significant, Z_s = - 0.0751, p value = 0.3759. Figure 11 indicated a visual representation of the general trend before and after the changepoint. The magnitude for the trend in both cases were 0.0306 and -0.0119 respectively, Table 6.

Table 6. MK test and Sen's Slope estimates for the two segments.

Segment	MK Trend test	Sen's-slope	P-value
1^{st} Segment	0.1570	0.0306	0.0240
2^{nd} segment.	-0.0751	-0.0119	0.3759

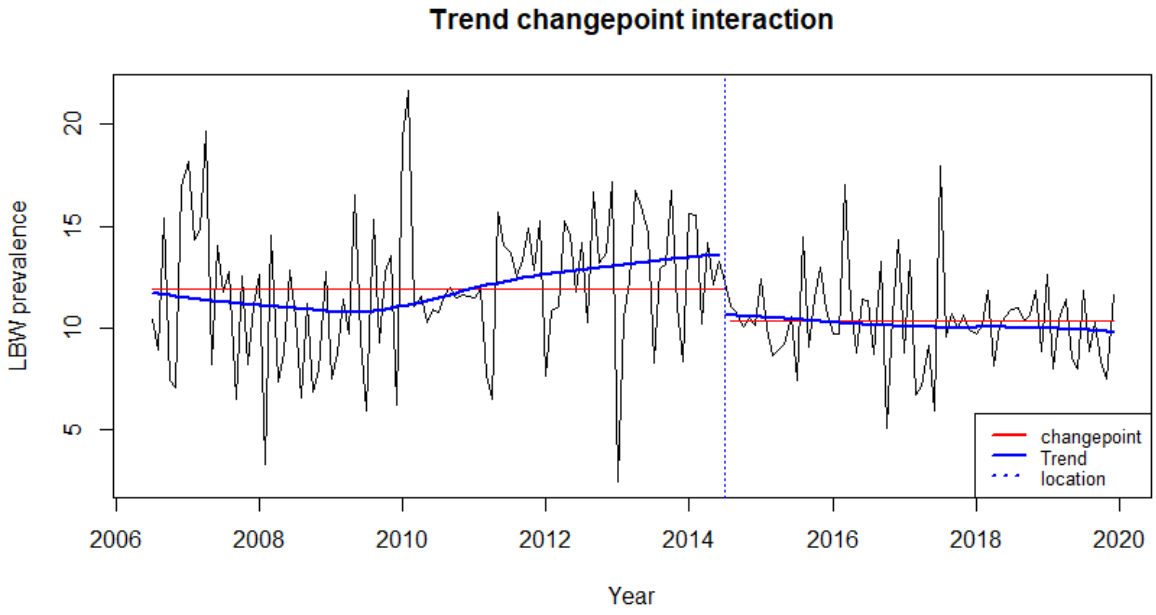

Figure 11. General trend from a fitted locally weighted least squares (lowess) regression before and after the changepoint.

4.7 Time series modelling.

The next sections discusses the fitting of multiplicative Seasonal Autoregressive Integrated Moving Average model that helped to study the pattern of LBW prevalence from the predicted series. The fitted model and the general trend of the predicted series can be seen in section 4.12.

4.8 Stationarity of the LBW prevalence data.

Figure 12, is the plot of the original series of LBW prevalence. Also shown in Figure 13 are the autocorrelation and partial autocorrelation plots of the original series. Time series models require that the data being used for modelling is stationary in its mean, variance and autocorrelation structure. Analysis of the ACF plot, Figure 13, reveals non-stationarity of the LBW observations. The observed spikes of the ACF plot depicts a geometric decay indicating the presence of a systematic trend in the data. This finding necessitated a first difference of the time series. Figure 14 show the resulting plot of the data after the first difference was passed to the data. Inspecting the resultant plot, we could remark a stabilized series. A further application of the ADF test to ascertain for the stationarity was passed to the data to ascertain the stationarity. The ADF test in this context test the null hypothesis that the first order LBW prevalence observations are not stationary. The results from the test gave; ADF= -8.035 and a P-value=0.01 which signified rejection of the null hypothesis.

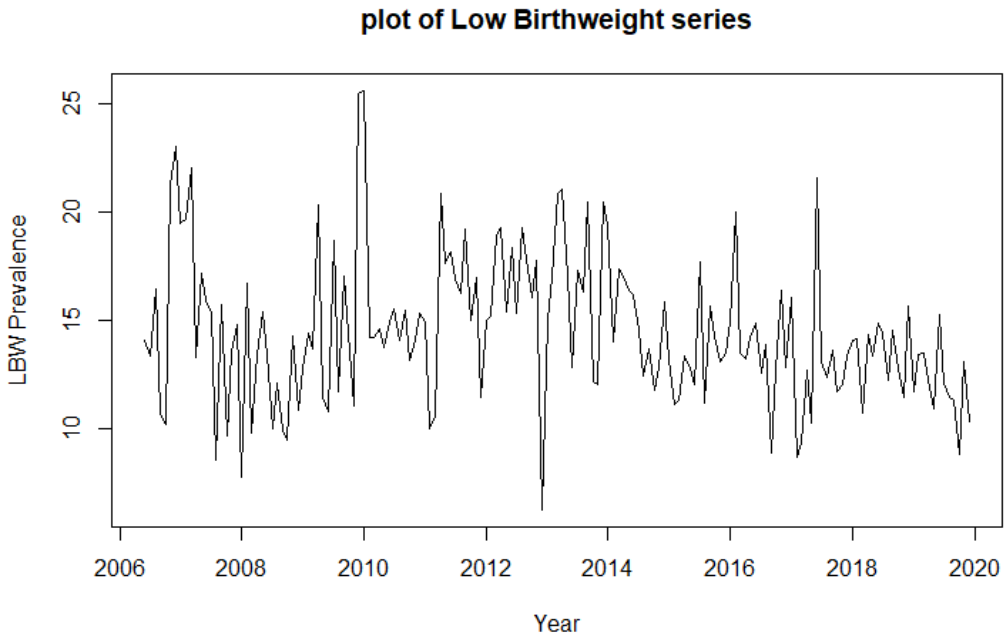

Figure 12. Plot of original low birthweight prevalence for the year 2006 to 2019.

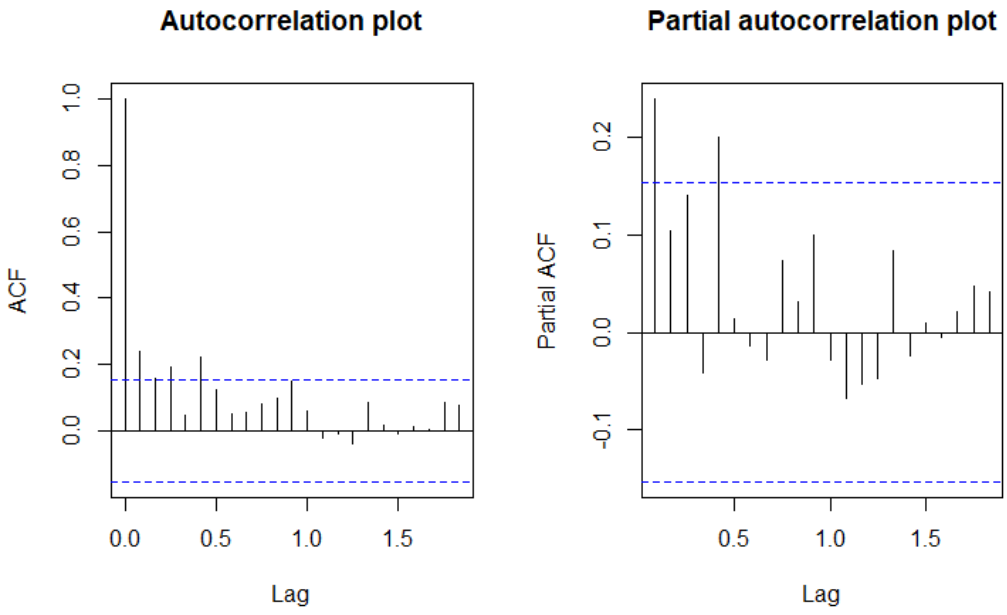

Figure 13. Auto-correlation and partial autocorrelation plots of the original series..

4.9 Model identification: Box-Jenkins Approach.

The approach provided by Box-Jenkins described in chapter 3, was followed in identifying the best model to fit the LBW prevalence data. After achieving stationarity in the data, we obtained p and q of AR and MA models respectively by analyzing the ACF and PACF plots of the stationary series. Below are the plots of the differenced series, Figure 15).

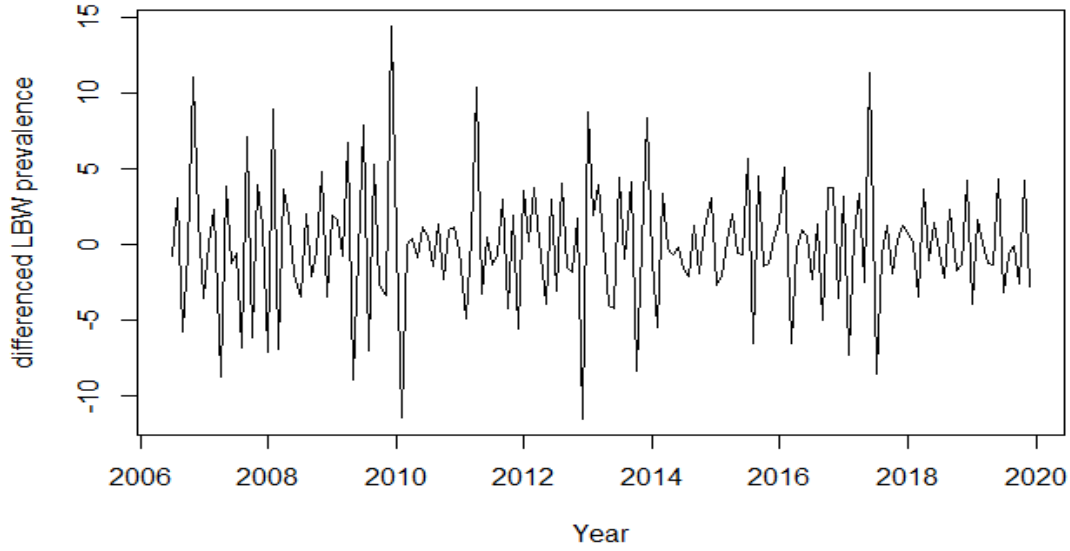

Figure 14. Differenced series of low birthweight prevalence

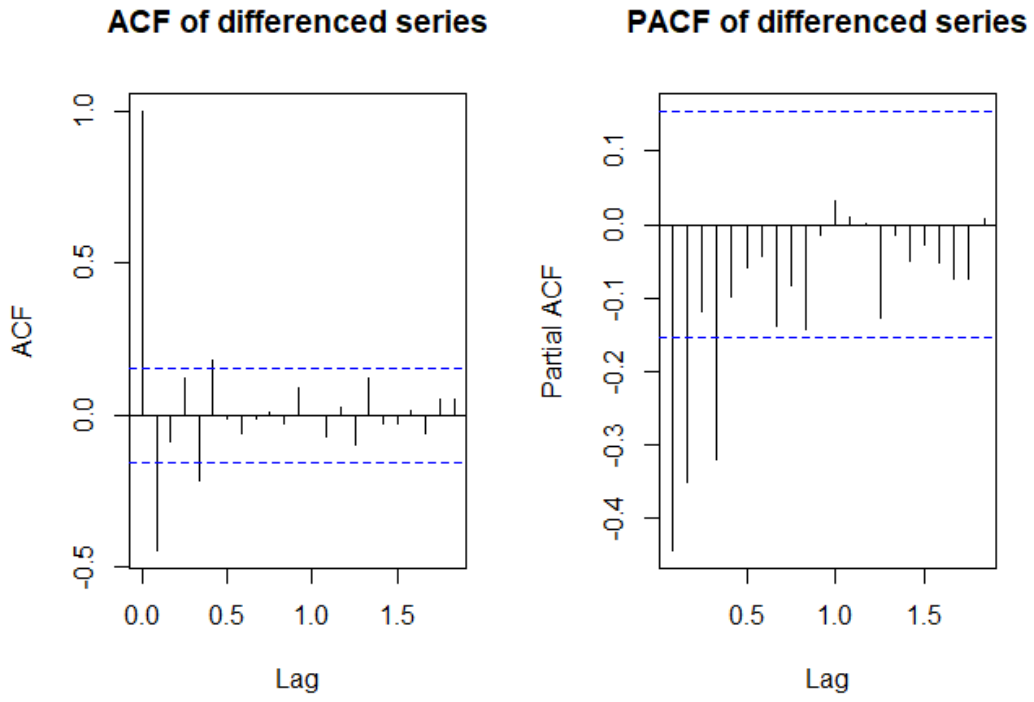

Figure 15. ACF and PACF plots of the differenced LBW prevalence series.

Inspecting both ACF and PACF graphs led to the following possible models to fit the data; ARIMA $(0,1,1)(0,1,1)_{12}$, ARIMA $(1,1,0)(0,1,1)_{12}$, ARIMA $(1,1,1)(1,1,1)_{12}$, ARIMA $(0,1,2)(0,1,1)_{12}$, ARIMA $(1,1,2)(0,1,1)_{12}$, ARIMA $(2,1,1)(0,1,1)_{12}$, ARIMA

$(2,1,2)(0,1,1)_{12}$ and ARIMA $(2,1,2)(1,1,1)_{12}$. The best chosen model for estimation had the least values in AIC, AICc and BIC as shown in Table 7.

Table 7. SARIMA models; AIC, AICc and BIC values.

Model	AIC	AICc	BIC
ARIMA $(0,1,1)(0,1,1)_{12}$	818.08	818.24	827.11
ARIMA $(1,1,0)(0,1,1)_{12}$	856.82	856.98	865.85
ARIMA $(1,1,1)(0,1,1)_{12}$	818.65	818.93	830.69
ARIMA $(1,1,1)(1,1,1)_{12}$	820.58	820.99	835.63
ARIMA $(0,1,2)(0,1,1)_{12}$	818.58	818.85	830.62
ARIMA $(1,1,2)(0,1,1)_{12}$	818.18	818.59	833.23
ARIMA $(2,1,1)(0,1,1)_{12}$	820.61	821.03	835.67
ARIMA $(2,1,2)(0,1,1)_{12}$	820.18	820.77	838.24
ARIMA $(2,1,2)(1,1,1)_{12}$	822.17	822.96	834.24

Based on the AIC, AICc and BIC values, the model with the least corresponding values was ARIMA $(0,1,1)(0,1,1)_{12}$ which was equivalent to the following form;

$$(1-B^m)^D(1-B)^d\, Z_t = \Psi_{\mathrm{Q}}(B^{\mathrm{m}})\rho(B)\varepsilon_{\mathrm{t}}$$

$$(1-B-B^{12}+B^{13})\, Z_t = (1+\rho B+\Psi B^{12}+\rho\Psi B^{13})\varepsilon_t \tag{39}$$

From the equation 39 a simplified model that can be used for prediction takes the form.

$$Z_t - Z_{(t-1)} - Z_{(t-12)} + Z_{(t-13)} = \varepsilon_t + \rho\varepsilon_{(t-1)} + \Psi\varepsilon_{(t-12)} + \Psi\rho\varepsilon_{(t-13)}$$

$$Z_t = Z_{(t-1)} + Z_{(t-12)} - Z_{(t-13)} + \varepsilon_t + \rho\varepsilon_{(t-1)} + \Psi\varepsilon_{(t-12)} + \Psi\rho\varepsilon_{(t-13)} \tag{40}$$

4.10 Parameter estimation.

Having identified the SARIMA model, we then estimated the model parameters. The method of maximum likelihood estimation was used in this the exercise. The estimated parameter values $MA_1 = -0.8590, SMA_1 = -0.9998$ were both significant, Table 8

Hence the resulting fitted model is of the form;

$$Z_t = Z_{(t-1)} + Z_{(t-12)} - Z_{(t-13)} + \varepsilon_t - 0.8590\ \varepsilon_{(t-1)} - 0.9998\ \varepsilon_{(t-12)} + 0.8588\ \varepsilon_{(t-13)} \tag{41}$$

Table 8. Parameter estimates

	Estimate	Standard error	p value
$MA1$	-0.8590	0.05659	2.2×10^{-16}
$SMA1$	-0.9998	0.1304	1.779×10^{-14}

4.11 Model Diagnostics.

In evaluating the validity of the model, residuals from the SARIMA fit were examined. Herein, residuals are expected to be independent with no systematic pattern. A plot of the model's residuals Figure 16 showed no possible systematic trends.

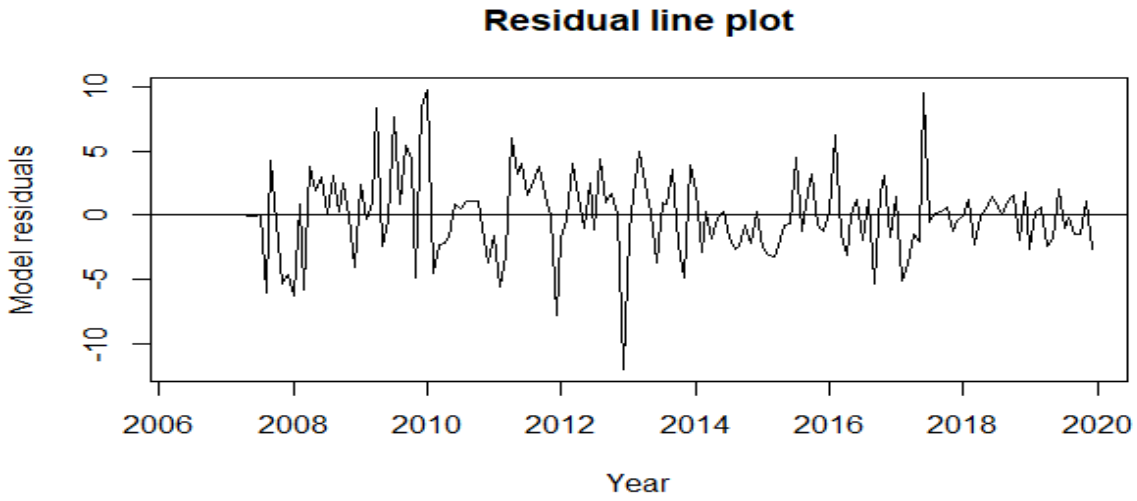

Figure 16. A plot of the SARIMA model residuals

To examine for any serial dependencies in the error terms, we plotted the ACF and PACF, Figure 17. Inspecting the plots, it was noted that all the spikes were not different from zero indicating absence of serial dependency in the residuals.

Ljung-Box statistic explained in chapter 3, was also applied to ascertain whether the model's residuals followed the white noise process. Implementing the test on R software yielded; Chi square= 0.9017, P-value= 0.3423. The null hypothesis of randomness was upheld indicating a parsimonious model.

4.12 Predicted trend using ARIMA $(0,1,1)(0,1,1)_{12}$.

From the estimated model 41, we predicted the prevalence of LBW in the next 24 months. The table 9 shows the estimated values for ARIMA (0,1,1) (0,1,1)12. A plot of the out-

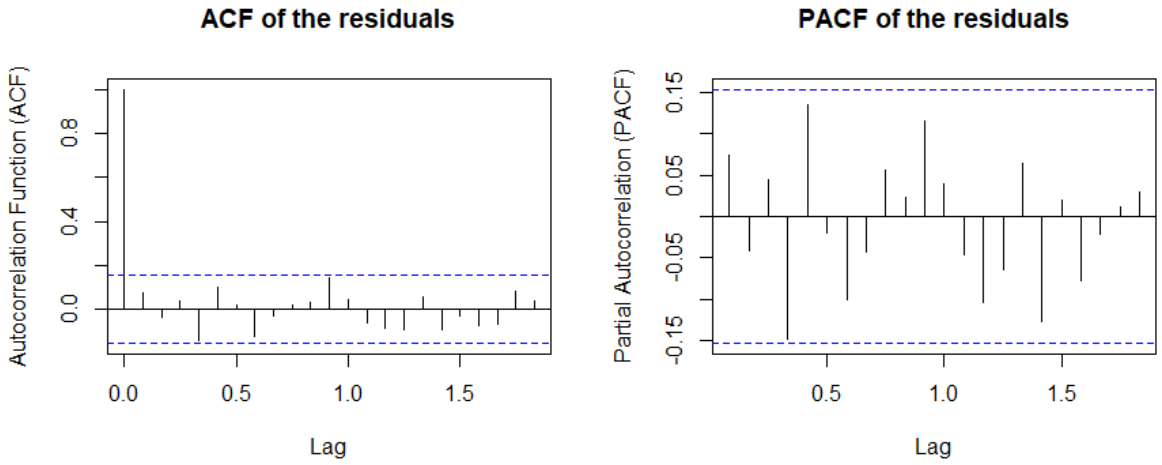

Figure 17. Autocorrelation and partial autocorrelation plots for the residuals.

sample values, figure 18 seemingly followed the declining trend in the original series.

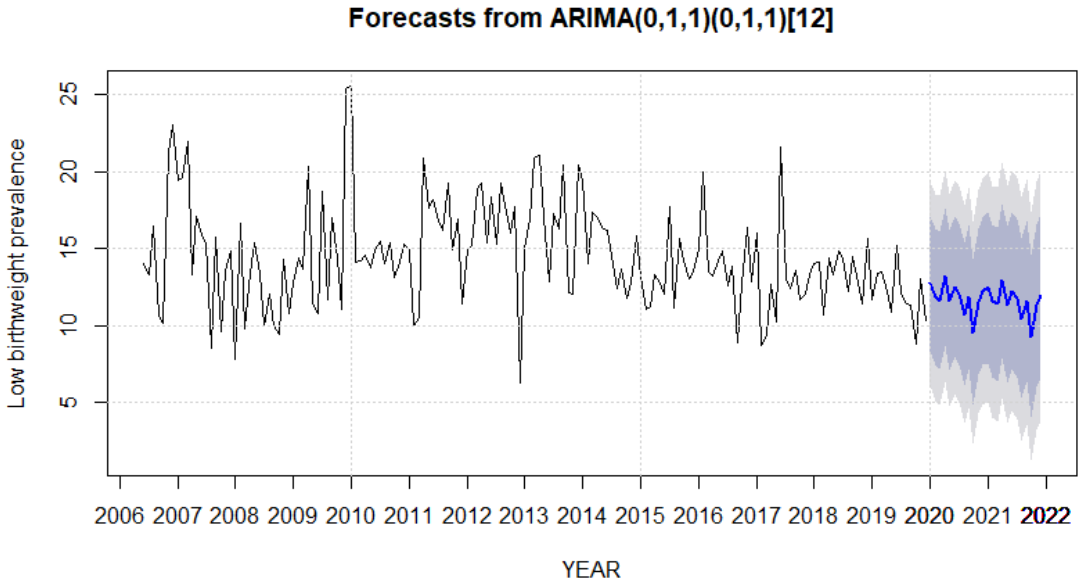

Figure 18. Estimated pattern using ARIMA $(0,1,1)(0,1,1)_{12}$ for the period of 2020 and 2021

Table 9. Forecasted values for ARIMA $(0,1,1)(0,1,1)_{12}$ model

YEAR 2020				YEAR 2021		
Month	**Estimate**	**95% CI**		**Estimate**	**95% CI**	
January	12.7250	6.0712	19.3789	12.4623	4.9920	19.9325
February	11.8440	5.1244	18.5636	11.5812	4.0439	19.1184
March	11.6320	4.8472	18.4167	11.3692	3.7656	18.9728
April	13.1809	6.3317	20.0302	12.9181	5.2487	20.5876
May	11.6283	4.7152	18.5414	11.3655	3.6309	19.1002
June	12.5009	5.5326	19.4692	12.2381	4.4423	20.0340
July	12.0463	5.0155	19.0772	11.7835	3.9243	19.6428
August	10.7174	3.6246	17.8101	10.4546	2.5325	18.3766
September	11.8456	4.6914	18.9997	11.5828	3.5984	19.5672
October	9.5567	2.3416	16.7718	9.2939	1.2477	17.3402
November	11.5440	4.2685	18.8194	11.2812	3.1736	19.3888
December	12.2010	4.8656	19.5363	11.9382	3.7696	20.1067

4.13 Seasonal variation of LBW prevalence.

As defined earlier, seasonality is the periodic pattern that reoccurs at intervals of time within the year. To evaluate the seasonal variations in LBW prevalence, we used the Moving Average method as explained in section 3.14 to obtain the seasonal indices associated with the months of the year.

Seasonal index in this context measures how much the average LBW prevalence of a particular month tends to be above or below the trend. Based on the indices obtained (figure 19), the month of April has the highest index (1.696) while the month of October had the lowest seasonal index (-1.688). This indicated that the prevalence of LBW in the month of April is expected to be above the seasons' trend value by 1.696 and reduce by 1.688 in the month of October. The months of January (0.663), April (1.696), June (0.942), July (0.544), September (0.696) and December (0.252) had increased prevalence from the expected rates, while the months of February (-0.303), March (-0.724), May (-0.300), August (-1.158), October (-1.688) and November (-0.620) had LBW prevalence below the trend values, Table 10

Notably, the seasonal index tends to rise above the expected values in the months of December and January after the short rains in the months of October and November.

The wet condition usually lead to an increased parasite density hence increased rates of malaria infection during that period. The increased LBW rates in the months of December and January meant that the mothers were in their last trimester when the parasite density increased during the short rainy season hence a negative impact on the baby's weight. During the long rains that are witnessed in April to June, a positive index implying increased LBW prevalence was recorded in April, June and also in the month succeeding the rainy season.

The lowest prevalence in the month of October can be attributed to the preceding dry season of July to September where the parasite density was expected to decrease due to unfavorable breeding grounds for malaria vectors.

Table 10. Monthly seasonal indices in prevalence of LBW. (Adj SV- Adjusted Seasonal variation)

Year	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec
2007							0.448	-5.754	2.022	-3.559	0.439	1.872
2008	-4.847	4.131	-2.684	1.241	3.131	1.344	-2.119	-0.173	-2.358	-3.286	1.417	-1.785
2009	-0.120	1.182	0.165	6.360	-2.610	-3.734	3.019	-4.493	0.797	-1.614	-4.810	9.317
2010	9.418	-1.952	-1.928	-1.444	-2.349	-0.923	0.543	-0.273	1.423	-0.958	-0.438	0.467
2011	-0.106	-5.160	-4.904	5.209	1.759	2.341	1.219	0.362	2.790	-1.733	0.384	-5.048
2012	-1.437	-1.268	2.400	2.744	-1.233	1.902	-0.917	3.002	1.357	-0.575	1.001	-10.313
2013	-1.417	0.561	4.489	4.743	1.134	-3.434	0.217	-0.810	3.585	-4.384	-4.293	3.974
2014	2.953	-2.179	1.600	1.586	0.921	0.857	-0.304	-2.058	-0.427	-1.946	-0.581	2.777
2015	0.187	-2.026	-1.862	-0.002	-0.703	-1.289	4.392	-2.596	1.442	-0.077	-1.324	-1.090
2016	0.480	5.656	-0.697	-0.599	0.313	0.814	-1.487	0.276	-4.070	-0.109	3.788	0.129
2017	3.006	-4.257	-3.770	-0.550	-2.779	8.707	0.251	-0.570	0.387	-1.659	-1.560	-0.103
2018	0.873	0.924	-2.528	1.013	-0.014	1.396	0.954	-1.113	1.103	-0.661	-1.774	2.521
2019	-1.345	0.440	0.723	-0.254	-1.477	3.009	-	-	-	-	-	-
SV	0.637	-0.329	-0.750	1.670	-0.326	0.916	0.518	-1.183	0.671	-1.713	-0.646	0.227
Adj SV	**0.663**	**-0.303**	**-0.724**	**1.696**	**-0.300**	**0.942**	**0.544**	**-1.158**	**0.696**	**-1.688**	**-0.620**	**0.252**

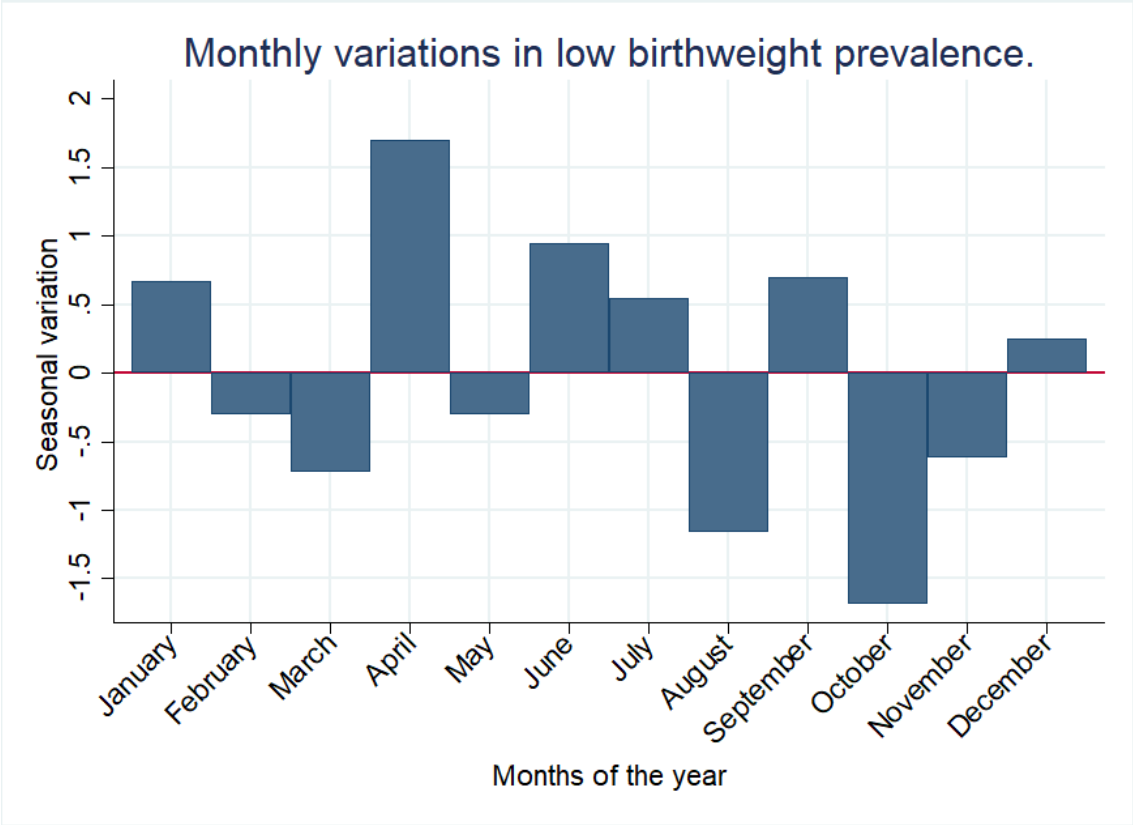

Figure 19. seasonal indices for LBW prevalence.

4.14 Malaria endemicity.

As explained earlier in chapter 3, Brabin et al. (1999) described the metrics of evaluating malaria transmission in Africa. Odds ratios of LBW risk among the primiparity group signified risk to malaria while the low birthweight prevalence was attributed to the excess risk from all other causes. In this section we considered using the odds ratios to examine the dynamics of malaria transmission.

Figure 20 shows the whole region stratified into KHDSS North and KHDSS South. Reflecting on the whole region, higher odds of LBW among the primiparity was observed in the year 2012 while the lowest was in 2008. In the year 2006 to 2008, the risk of LBW was observed to decrease from (OR=1.73; CI 1.105 - 2.697 to OR=1.17; CI 0.847 - 1.634). A slight increase was recorded in 2009 which later decreased in 2010. The OR begun to rise in 2011 (OR=1.19; CI 0.918 - 1.542) to 2012 (OR=1.92; CI 1.509 - 2.444). In 2013 a drop was observed which rose in 2014. Thereafter, a decreasing trend in 2014 (OR=1.781; CI 1.449 - 2.19) to 2017 (OR=1.227 CI 0.871 - 1.728) was remarked. It was during this period when the change point had been identified in prevalence of LBW,(see 4.6). A rising risk was later observed in 2018 and 2019.

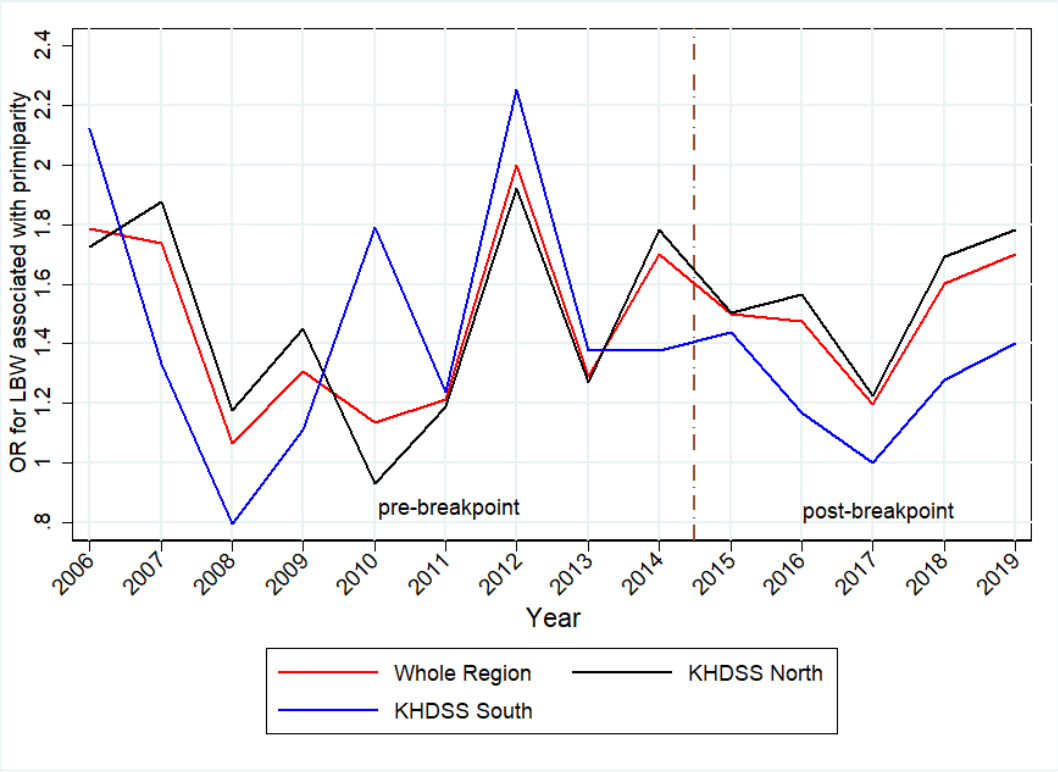

Figure 20. Changing OR for LBW associated to primiparity with reference to multiparity since 2006-2019.

In KHDSS North and KHDSS South, odds ratios seemingly assumed the same pattern. The risk of LBW was higher in 2012 for both regions; Kilifi HDSS South (OR=2.3) , Kilifi HDSS North (OR=1.9). In KHDSS South, the risk of LBW was observed to decline from 2006 to 2008 and later rose to 2012 upon which a decline was remarked thereafter. In KHDSS North, the declining trend was seen to 2010 and later rose to 2012 after which a downward trend in the risk of LBW was observed towards 2017. In both regions the risk begun to rise from 2018 to 2019.

Turning the focus to how LBW exhibit itself in the different months of the year, figure 21 , we could observe the lowest OR in February and quite higher in the month of June. The month of February is characterized as among the dry seasons of the year while the month of June is among the wet seasons.

Wet season in the region is experienced from the month of April through June and later in the month of October and November while the rest of the months are characterized as dry season. During the wet season the trend was observed to exhibit an upward trajectory; (April, OR = 1.38, May, OR = 1.49 and June, OR = 1.80) and (October, OR = 1.50, November, OR = 1.58).

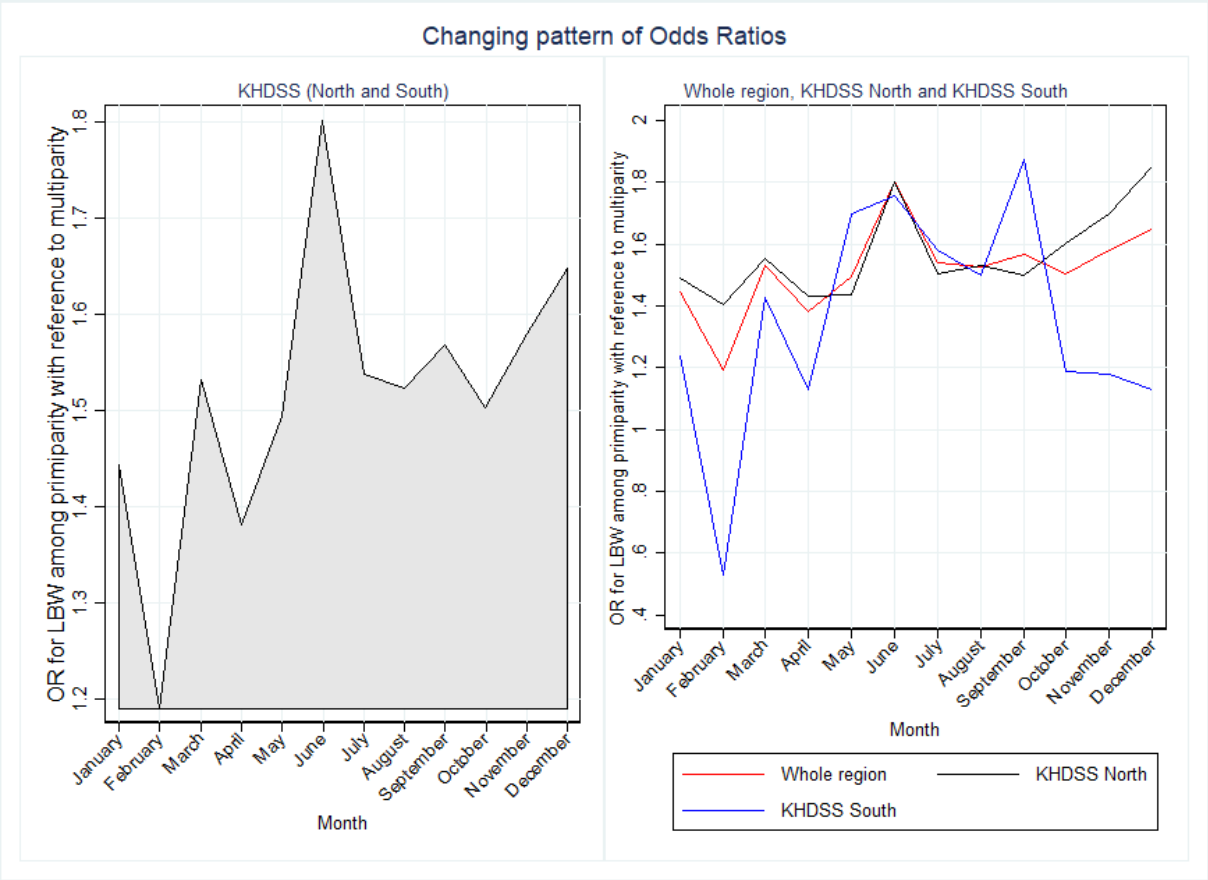

Figure 21. Changing risk of LBW among the primiparity in different months of the year.

Shortly after the long rainy season in June, the trend is observed to decline. The months of January (OR=1.44) and February (OR=1.19) also recorded low values except for the month of December (OR=1.65) which marks the beginning of the dry season after the short rains of October and November.

Table 11 shows the LBW prevalence of both parities and OR for LBW in babies born to primiparity group for 16 locations in the two regions of Kilifi HDSS North and Kilifi HDSS South,

Table 11. LBW prevalence of both primiparity and multiparity, OR for LBW in Primiparity with reference to Multiparity for the period of 2006-2019

		LBW prevalence	Odds Ratio		
Location	**Parity**	**Number (%)**	**OR**	**95 % CI**	
Kilifi HDSS North					
Kilifi Township	Primiparity	818 (15.20)	1.486	1.343	1.644
	multiparity	924 (10.77)			
Matsangoni	Primiparity	89 (22.08)	1.343	0.976	1.849
	multiparity	100 (17.39)			
Ngerenya	Primiparity	138 (18.80)	1.889	1.447	2.446
	multiparity	125 (11.00)			
Roka	Primiparity	95 (15.83)	1.017	0.763	1.351
	Multiparity	139 (15.64)			
Gede	Primiparity	6 (14.29)	1.314	0.328	5.271
	multiparity	4 (12.50)			
Sokoke	Primiparity	52 (17.87)	1.414	0.919	2.130
	multiparity	57 (13.35)			
Takaungu / Mavueni	Primiparity	259 (19.65)	1.820	1.509	2.195
	multiparity	262 (11.86)			
Tezo	Primiparity	379 (17.11)	1.6656	1.432	1.937
	multiparity	405 (10.95)			
Kilifi HDSS South					
Banda ra Salam	Primiparity	40 (25.64)	1.468	0.895	2.408
	multiparity	42 (18.83)			
Chasimba	Primiparity	71 (18.25)	1.306	0.924	1.846
	multiparity	84 (14.61)			
Jaribuni	Primiparity	15 (17.05)	1.758	0.792	3.902
	multiparity	14 (10.69)			
Junju	Primiparity	198 (21.45)	1.404	1.125	1.753
	multiparity	188 (16.26)			
Kauma	Primiparity	33 (18.97)	1.278	0.770	2.120
	multiparity	41 (15.19)			
Mtwapa	Primiparity	120 (18.40)	1.091	0.834	1.428
	multiparity	141 (17.13)			
Ziani	Primiparity	54 (19.01)	1.495	1.027	2.255
	multiparity	69 (13.77)			

The prevalence of LBW among the primiparity ranged from 14.29 % in Gede location to 25.64 % in Banda ra Salam location, while among the multiparity group the prevalence ranged from 10.69 % in Jaribuni location to 18.83 % in Banda ra Salam. The odds ratio for LBW prevalence in the primiparity group in comparison to multiparity group ranged from 1.014 in Roka location to 1.876 in Ngerenya location. Considering the OR value of 1.7, Brabin et al. (1999) as the cut-off point to distinguish between malaria exposed populations, it was noted that three locations had an OR > 1.7. these included Jaribuni location (OR=1.758), Ngerenya location (OR=1.889), and Takaungu/Mavueni location (OR=1.820)

It was further observed that, in an area suggestive of malaria transmission, (OR > 1.7) the proportion of multiparity group having LBW babies was notably low while those of primiparity were high; Jaribuni location (OR=1.758, LBW prevalence=10.69 %, 17.05 %), Ngerenya location (OR=1.889, LBW prevalence = 11.00 %, 18.80 %) and Takaungu/Mavueni location (OR=1.820, LBW prevalence =11.86 %, 19.65 %), (Herein, the order of LBW prevalence are as multiparity and primiparity respectively).

Moreover, high proportions of multiparity group having LBW babies were observed in some areas indicative of low transmission, (OR < 1.7). Some of these locations included Junju (aOR = 1.404, LBW prevalence = 16.26), Matsangoni (aOR = 1.343, LBW prevalence = 17.39 %) and Mtwapa (OR =1.091, LBW prevalence = 17.13).

In KHDSS North, odds of LBW was higher among the Primiparity as comapared to multiparity in Takaungu (OR=1.82) and Ngerenya (OR=1.88) locations while it was lower in Roka location (OR=1.017). In KHDSS South, it was higher in Jaribuni (OR=1.758) while lower in Mtwapa location (OR=1.091) .

4.15 Trend in Low Birthweight ratios of Primiparity in relation to Multiparity.

In this section we analyze the trends in bi-monthly LBW ratios of primiparity in relation to multiparity groups. (Bi-monthly ratios were chosen over the monthly ratios because some months had zero counts in either the primiparity or the multiparity category hence yielding zero or infinite ratios.)

Figure 22 indicate the pattern of LBW ratios in different regions. A significant trend in the ratios was observed in the whole region (Both South and North combined), p value < 0.05. In KHDSS North and KHDSS South Mann-Kendall test indicated an insignificant upward trajectory, p value > 0.05, Table 12.

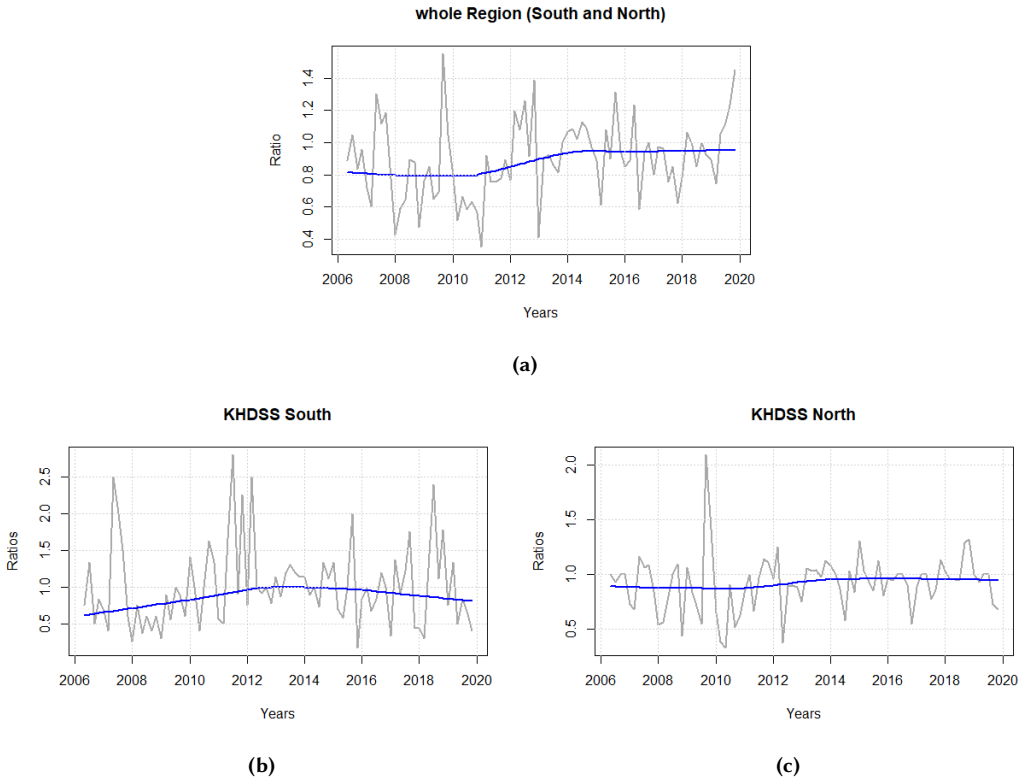

Figure 22. Low Birthweight ratios of primiparity to multiparity in (a) Whole Region - KHDSS North and South, (b) KHDSS South and (c) KHDSS South

Table 12. Trend significance of LBW ratios of primiparity in relation to multiparity.

Region	Mann-Kendall	Sen's slope	P-Value
Whole Region	0.0674	0.0027	0.015
Kilifi HDSS North	0.080	0.001	0.297
Kilifi HDSS South	0.0674	0.002	0.378

4.16 Spatial distribution of Low birthweight cases.

In this section the analysis was a two step process: First, we mapped the LBW prevalence in the geographical area with the aim of getting a visual representation of the spatial pattern in the data. Secondly, we used the Spatial Scan statistic to identify the local clusters of low birthweight in the study area.

4.16.1 Mapping of Low Birthweight prevalence.

For the period of 2006-2019, Figure 23 An eyeballing indicated that a most likely cluster of locations in the Southern region had high rates of LBW over time. These included Mtwapa, Junju, ziani, Banda ra Salam, Chasimba and Kauma. In general, the trend of LBW prevalence decreased from the period of 2006-2012 to the period of 2013-2019. In the most recent year, (2019), a mostly likely cluster of locations with high rates was observed in the southern region; Chasimba, Ziani, Banda ra Salam and Junju while a possible cluster of locations with lower rates were observed in the Northern region.

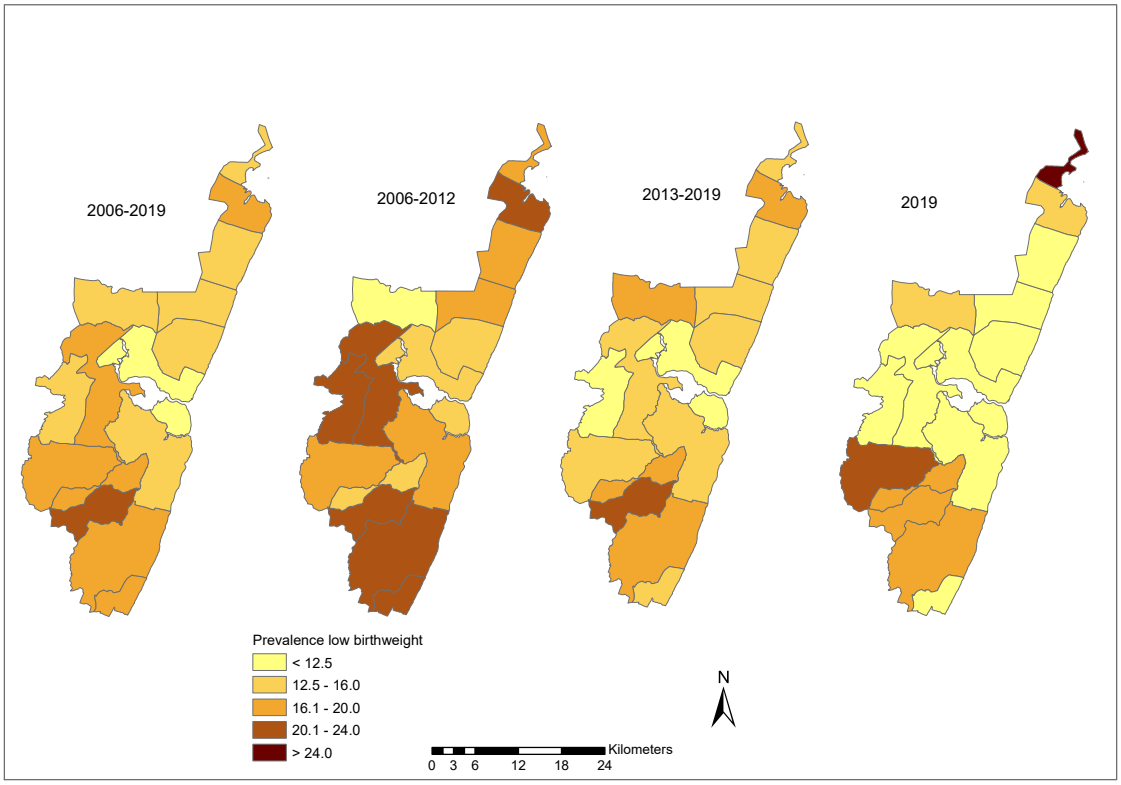

Figure 23. Spatial distribution Low birthweight prevalence in KHDSS

4.16.2 Spatial Scan Statistic results

In the year 2019, SaTScan detected one significant cluster/hotspot in the southern region (log likelihood ratio = 13.59, p value = 0.003), figure 24 . The cluster was centered at 3.793550 S, 39.703766 E and had a radius of 9.71 Km. It comprised of Junju, Banda ra salam, Ziani and Chasimba locations. These locations compared well with the spatial pattern described in figure 23 . The other identified clusters were insignificant, p value > 0.05.

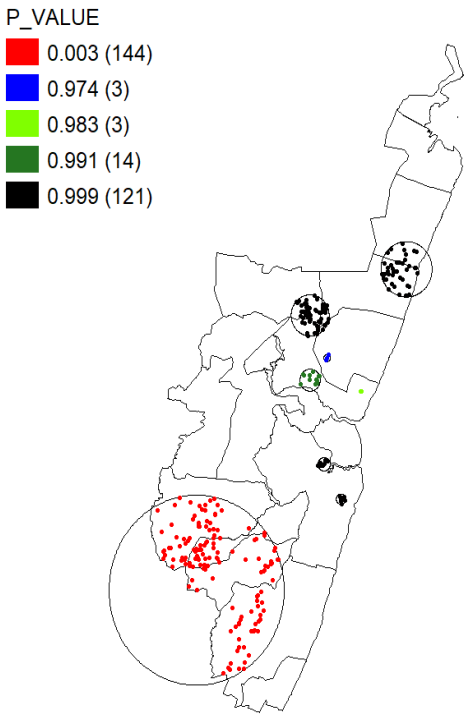

Figure 24. Local clusters of high rates identified by SaTScan for the year 2019

Considering a three year period interval from 2006 then 2009,2012,2015 and 2018, Table 13, SaTScan identified five clusters of high rates in 2006, out of which one cluster centered at 3.612337 S, 39.753371 E in Kilifi township was significant (log likelihood ratio = 5.58, p value = 0.038). In 2009, two clusters with high rates were identified both in the northern and southern part of KHDSS. A significant one centered at 3.8255785 S, 39.7764419 E was located in the Southern region (log likelihood ratio = 17.43, p value = 0.0001) covering Ziani, Chasimba, Takaungu, Mtwapa, junju and Banda ra Salam locations. In 2012, Roka, Matsangoni and Gede locations in the Northern region formed a significant hotspot centered at 3.3379805 S, 39.937786 E (log likelihood ratio = 7.91, p value = 0.014). Both in 2015 and 2018 two clusters were identified, Table13 but non of them was significant.

Table 13. Cluster of locations (and their significance) detected by Spatial Scan Statistic. Monte Carlo simulation with 999 repetitions was used to evaluate the significance of the identified clusters.

Year	Cluster	Radius (Km)	Locations	Observed cases	Expected cases	Relative Risk	Likelihood Ratio	p value
2006	1	6.4	Kilifi Township	11	4.02	2.9	5.583	0.038
	2	1.37	Mtwapa	7	3.24	2.22	2.123	0.527
	3	2.98	Chsimba	6	3.4	1.8	1.03	0.904
	4	2.94	Tezo	22	18.08	1.26	0.560	0.991
	5	3.4	Banda ra salam	3	1.77	1.78	27.3	0.995
2009	1	13.08	Banda ra salam,junju, chasimba, takaungu, Ziani,Mtwapa	98	58.77	2.01	17.43	0.0001
	2	4.22	Ngerenya	10	6.73	1.5	0.857	0.999
2012	1	13.96	Roka, Gede, Matsangoni	34	17.14	2.00	7.91	0.014
	2	9.65	Mtwapa,junju, Ziani Banda ra salam	54	37.19	1.52	4.60	0.254
2015	1	10.77	junju,Takaungu, ziani Jaribuni, Chasimba Kauma,Banda ra salam	127	100.67	1.36	4.815	0.182
	2	3.19	Sokoke	4	1.59	2.54	1.599	0.989
2018	1	4.85	Roka, matsangoni	24	14.8	1.64	2.89	0.82
	2	3.3	Chasimba	17	9.87	1.74	2.59	0.885
2019	1	9.71	Junju,Chasimba, Ziani, Banda ra salam	60	30.92	2.05	13.59	0.0027
	2	0.37	Tezo	3	0.36	8.32	6.35	0.974
	3	1.12	Kilifi township	7	1.81	3.90	5.49	0.991
	4	0.52	Takaungu	4	0.72	5.56	4.90	0.999
	5	0.68	takaungu Mavueni	5	1.21	4.17	4.30	0.999
	6	2.84	Ngerenya, Roka	19	9.96	2.00	4.250	0.999
	7	2.11	Sokoke,Kilifi Township Ngerenya, Tezo	16	7.61	2.14	2.24	0.999

5 Discussions, Conclusion, Limitations and Recommendations

5.1 Overview.

The primary objective of the research was to evaluate trends in LBW over time so as to explain the malaria transmission. This chapter thus presents a discussion on the results obtained from the previous chapter. Thereafter, conclusions, limitations and recommendations are documented.

5.2 Discussion.

5.2.1 Low birthweight prevalence.

LBW prevalence in the region was 14.23%, a statistic which was higher than the national prevalence (11.5%) and slightly below the global estimate of 15.5%. Other studies in the country include a study in Olkalou District Hospital which reported a prevalence of 12.3%, (Muchemi et al., 2015). Our obtained prevalence was lower than 15.0% prevalence recorded in a study in Nyanza provisional Hospital, (Were et al., 2002). Primigravids registered a higher prevalence when compared to multigravida. The higher prevalence among the primigravids was also in agreement with other studies, (Mutabingwa et al., 2005). The higher prevalence can be attributed to their high susceptibility to malaria infection, BRABIN (1991). The two regions; Kilifi HDSS North and Kilifi HDSS South also had a significantly different prevalence with South having a higher prevalence. That implied that emphasis on LBW control mechanisms needed to be done on the southern part of the region.

5.2.2 Trends and seasonal variation of low birthweight.

Low birthweight prevalence in both parities was observed to decline steadily from 2006 to 2008. Citing references from other studies done in the same area, Mogeni et al. (2016) in their study had also reported a declining malaria positive fraction from mid-1990 to almost zero by 2008, thus this could explain the drastic decrease in the LBW prevalence. Increasing trend was later observed upto 2012 upon which a decline was noted.

In general, we found a significantly decreasing trend in LBW prevalence. On further analysis using change-point techniques, we could realize a significant change-point at

mid 2014. The low birthweight series before the change-point indicated a significant increasing trend while the series after the changepoint indicated a decreasing insignificant trend. Hence relying only on the general trend over time could be misleading. Between July, 2012 and October, 2015, the government's mass roll-out of bed nets aimed at achieving universal coverage, Kamau et al. (2017) may have had a substantial impact on malaria prevention, hence the beginning of the downturn in mid-2014.

LBW prevalence among the primiparity was higher throughout the study period when compared to that of multiparity. The pattern of the curve had similar properties to that of the overall prevalence. The two regions (North and South of the KHDSS) showed significant differences on the LBW prevalence. The Southern region of the HDSS depicted a higher prevalence throughout the study period. Drawing much from (Mogeni et al., 2016), the higher prevalence in the southern region could be explained by the findings made by the authors that malaria parasites were consistently higher in the Southern region as compared to the Northern region. In addition, the low prevalence in the Northern region could be explained by the rising prevalence on the use of ITN which by 2009 and 2013 it had risen to 55.9% and 82.6% respectively.

Low birthweight also varied with the months of the year. The seasonal variation of LBW could be explained by the bimodal rainfall pattern experienced in the area. The region experiences a period of long rainfall in April through June and short period of rains in October and November. Positive seasonal indices ware recorded in the months of April, June and also in the month preceding the rainy season, July, with the month of April registering the highest index. The positive variations could be attributed to the increased parasite density due to favorable climatic conditions for mosquito breeding which lead to increased rate of malaria infection and a significant decrease in birthweights. The lowest prevalence was recorded in the month of October which could be attributed to low transmission of malaria due to the preceding dry season of July to September. The dry season is unfavorable for mosquito breeding hence reduced rate of infection. The dry months of December and January also had indices which were greater than the expected value which implied increased rates of LBW. The increased rates could be perceived to mean that, deliveries in the two months (December and January) were in their last trimester during the period of short rains experienced in the months of October and November which favored the multiplication of malaria vectors hence increased infection. Thus, the effect of LBW was felt shortly during those two months. This situation was also pointed out in a study by Cottrell et al. (2007) where they indicated that infection at the last stage of the pregnancy was a risk factor for LBW.

On modeling the LBW prevalence using the SARIMA model we could identify SARIMA $(0,1,1)(0,1,1)_{12}$ had the minimum AIC and BIC values hence fitted the data well. We then used the model to estimate the pattern of the prevalence in the next 24 months. It

was clearly noticeable that the trend was expected to follow the downward pattern in the original series.

5.2.3 Endemicity of malaria.

In this section, the aim was to evaluate malaria transmission using Low Birthweight. The metrics used are as explained by (B. J. Brabin et al., 1999). Malaria risk was assessed using the OR for LBW among the primiparity with reference to multiparity.

The odds ratios for LBW among the primiparity in KHDSS North and South were observed to seemingly assume the same pattern in different years except in 2010 where an increase was noticeable in South while in the North a decline was noted. In both regions a decline from 2006 to 2008 was remarked. It then rose slightly to 2009 and dropped in 2010, though the drop was only in the north and also in the whole region. The drop in 2006-2008 and then in 2010 was also in agreement with a studies done by (Macharia et al., 2018) and (Snow et al., 2015) where they had reported significant decrease in *P. falciparum* parasites prevalence during this period. (O'Meara et al., 2008) did also described a declining incidence in hospital admissions for malaria from 2004-2007. This could further be explained by the District Health Management free Insecticide Treated Nets campaign of September 2006 that saw the coverage increase from 0.25 to 0.5 ITN per person across Kilifi District, Okiro et al. (2007). From 2011, an up and down trajectory with increasing trend was observed until 2014 when a noticeable decline was remarked all through to 2017. The same scenario was the case in Macharia et al. (2018) when a decline in *pf* PR was observed after 2014. An increasing risk was then seen in the last two years of 2018 and 2019.

Malaria infections are influenced by the changing patterns of rainfall. In Kenya transmission is expected to increase during the period of long rains(April to June) and during the period of short downfall of rain in October through November. Other months are characterized as dry season where transmission intensity is expected to fall. In the two seasons, we observed the odds ratios to have an increasing trend shortly during the period of long rainfall, after which a decline was observed from the month of July. Later on, during the onset of the short rains in October, the trend was observed to rise again to the month of December after which a decline was realized. In general the higher odds of LBW among primiparity in relation to multiparity was experienced in June while the lowest was in February.

The adjusted OR for LBW prevalence ranged from 1.017 in Roka location to 1.889 in Ngerenya location. It was noted that locations suggestive of malaria transmission (i.e. with primigravids having an odds Ratio > 1.7 for LBW) were Jaribuni, Ngerenya and Takaungu/Mavueni. The remaining locations had an OR $\leq$ 0.7. It was further noted that, the prevalence of LBW among multiparity was low in locations indicative of malaria

transmission (OR > 1.7) while the prevalence among the primiparity in the same locations was a bit high. This phenomenon indicated that the multiparity group had developed a significant immunity against malaria infection while the primiparity's immune system was yet to develop hence rendered vulnerable to infection. The same scenario was reported by (Rogerson et al., 2007) where multigravids developed immunity with successive infected pregnancies.

High proportions of Multiparity who delivered LBW babies were observed in locations suggestive of low malaria transmission, (OR < 1.7). These locations included Junju, Matsangoni and Mtwapa. The higher proportions was explained by (B. J. Brabin et al., 1999) as a situation that was the case in low malaria transmission areas since the group in question was yet to develop the immunity against malaria.

5.2.4 Low Birthweight Ratio of Primiparity in relation to Multiparity.

As discussed in the previous section , malaria transmission was estimated by studying the relationship between OR for LBW among the primiparous group with reference to the multiparity. In this section we assessed the trend in LBW ratios of primiparity to multiparity. A significant increasing trend in the odds of LBW among the primiparity was recorded in the whole region, p value < 0.05. Both the two regions of KHDSS had a relatively increasing trend in the ratios but was not statistically significant, p value > 0.05.

5.2.5 Mapping of low birthweight cases.

Mapping of significant clusters with high rates of LBW prevalence was aimed at informing the health stakeholders on the specific geographical locations that needed a prioritized course of action. Based on the most recent study period (2019), a significant hotspot on the southern region of KHDSS was identified and this study recommended for preventive measures to the affected locations.

The presence of clusters identified by the SaTScan indicated a spatial heterogeneity of LBW prevalence in the region. Identification of these clusters provides information of the geographical areas that needed health interventions to cab the elevated risk of LBW.

In 2019 a significant pure spatial cluster covering Junju, Chasimba, Ziani and Banda ra Salam locations was identified. In 2006 a significant cluster was identified at Kilifi Township. Three years later a highly significant cluster in the southern region was identified which covered Ziani, Banda ra Salam Junju, Chasimba, Takauungu and Mtwapa. In 2012 Significant cluster was located at Roka, Gede and Maatsangoni locations.

5.3 Conclusion.

The research was directed primarily at testing whether odds ratio for LBW among the primiparity with reference to multiparity could serve as a cheap and inexpensive tool to define the transition of malaria risk. In that regard, this study has led to the understanding of LBW prevalence patterns over time. Findings offer information on the prevalent trend of LBW to health professionals and also to organizations tracking the progress of achieving the objective of reducing the prevalence of LBW by 30% by 2025. The general prevalence of LBW in the region was detected to increase significantly between 2006 and mid-2914 after which an insignificant decline was observed. A decreasing trend of LBW prevalence was expected in the next 24 months beginning 2020.

The OR for LBW among the first pregnancies were in comparison with the changing patterns of malaria transmission in different time periods and also with the patterns influenced by either the dry or wet seasons of the year.

Hence OR for LBW in first pregnancies was suggestive of the prevalence of plasmodium parasites, thus proposed as a cheap surveillance mechanism for malaria transition.

5.4 Limitations and caveats.

Part of the data collection process for May 2010 to December 2010 was affected by the COVID-19 pandemic. Other unforeseen events included the Nurses strike from July 2017 to October 2017 which affected the smooth flow of the maternity services.

The analysis has utilized the multiplicative seasonal Auto-regressive Moving Average model which assumes a Gaussian distribution of the data. The seasonal non-stationary Low birthweight data with low or zero count in some months could lead to inaccurate predictions from the transformed data. An alternative to multiplicative SARIMA model is Generalized Multiplicative Seasonal Autoregressive Integrated Moving Average (GSARIMA) model for Poisson and binomial distributed data.

Gestational age is one of the determinants of low birthweight. Our analysis lacked information on this main predictor which could have helped understand the context of LBW for every woman.

5.5 Recommendations.

Findings from this research suggested the following measures;
First, in order to realize a continually decreasing trend in the prevalence of LBW, the study encouraged all expectant mothers to seek for antenatal care services during their period of pregnancy.

Secondly, prevalence of LBW was constantly high among the primiparity group, a condition which has been alluded to their susceptibility to malaria infection due to underlying impaired immune system triggered by the pregnancy. The study recommends to the health stakeholders to consider maximizing on community awareness to increase knowledge on the available programs aimed at alleviating the deleterious effects of malaria infection.

Thirdly, KHDSS North and KHDSS South showed a varying LBW prevalence which signified higher rate of malaria infection in South than North. The study recommended that health interventions aimed at controlling malaria transmission be implemented in the whole region with emphasis on the southern region.

Finally, as a cheap and simple tool to monitor the risk of malaria in a population, the study proposes the use of odds ratios of LBW in primiparity with reference to multiparity.

5.6 Future Research

The current research was carried out in a region where malaria transmission is considered to be stable. Areas of lower transmission may be considered by future research. Moreover, the region under review was comparatively small, so the population need to be expanded.

Bibliography

Autino, B., Noris, A., Russo, R., and Castelli, F. (2012). Epidemiology of malaria in endemic areas. *Mediterranean journal of hematology and infectious diseases*, 4(1).

Bahizire, E., Dramaix, M., Bigirinama, R., Balegamire, S., Balungu, Y., Meuris, S., D'Alessandro, U., and Donnen, P. (2018). Prevention against malaria before the first antenatal visit and absence of anaemia at the first visit were protective from low birth weight: results from a south kivu cohort, democratic republic of the congo. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 112(8):383–392.

Beier, J. C., Killeen, G. F., and Githure, J. I. (1999). Entomologic inoculation rates and plasmodium falciparum malaria prevalence in africa. *The American journal of tropical medicine and hygiene*, 61(1):109–113.

Blanc, A. K. and Wardlaw, T. (2005). Monitoring low birth weight: an evaluation of international estimates and an updated estimation procedure. *Bulletin of the World Health Organization*, 83:178–185d.

Blencowe, H., Krasevec, J., de Onis, M., Black, R. E., An, X., Stevens, G. A., Borghi, E., Hayashi, C., Estevez, D., Cegolon, L., et al. (2019). National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *The Lancet Global Health*, 7(7):e849–e860.

BRABIN, B. (1991). An assessment of low birthweight risk in primiparae as an indicator of malaria control in pregnancy. *International journal of epidemiology*, 20(1):276–283.

Brabin, B., Agbaje, S., Ahmed, Y., and Briggs, N. (1999). A birthweight nomogram for africa, as a malaria-control indicator. *Annals of Tropical Medicine & Parasitology*, 93(sup1):S43–S57.

Brabin, B. J. (1983). An analysis of malaria in pregnancy in africa. *Bulletin of the World Health Organization*, 61(6):1005.

Brabin, B. J. and Rogerson, S. J. (2001). The epidemiology and outcomes of maternal malaria. In *Malaria in Pregnancy*, pages 44–71. CRC Press.

Cochrane, D. and Orcutt, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American statistical association*, 44(245):32–61.

Cot, M., Le Hesran, J. Y., Staalsoe, T., Fievet, N., Hviid, L., and Deloron, P. (2003). Maternally transmitted antibodies to pregnancy-associated variant antigens on the surface of erythrocytes infected with plasmodium falciparum: relation to child susceptibility to malaria. *American journal of epidemiology*, 157(3):203–209.

Cottrell, G., Mary, J.-Y., Barro, D., and Cot, M. (2007). The importance of the period of malarial infection during pregnancy on birth weight in tropical africa. *The American journal of tropical medicine and hygiene*, 76(5):849–854.

Cottrell, G., Moussiliou, A., Luty, A. J., Cot, M., Fievet, N., Massougbodji, A., Deloron, P., and Tuikue Ndam, N. (2015). Submicroscopic plasmodium falciparum infections are associated with maternal anemia, premature births, and low birth weight. *Clinical Infectious Diseases*, 60(10):1481–1488.

De Beaudrap, P., Turyakira, E., Nabasumba, C., Tumwebaze, B., Piola, P., Boum II, Y., and McGready, R. (2016). Timing of malaria in pregnancy and impact on infant growth and morbidity: a cohort study in uganda. *Malaria journal*, 15(1):92.

Desai, M., Gutman, J., Taylor, S. M., Wiegand, R. E., Khairallah, C., Kayentao, K., Ouma, P., Coulibaly, S. O., Kalilani, L., Mace, K. E., et al. (2016). Impact of sulfadoxine-pyrimethamine resistance on effectiveness of intermittent preventive therapy for malaria in pregnancy at clearing infections and preventing low birth weight. *Clinical Infectious Diseases*, 62(3):323–333.

Desai, M., Ter Kuile, F. O., Nosten, F., McGready, R., Asamoa, K., Brabin, B., and Newman, R. D. (2007). Epidemiology and burden of malaria in pregnancy. *The Lancet infectious diseases*, 7(2):93–104.

Eisele, T. P., Larsen, D., and Steketee, R. W. (2010). Protective efficacy of interventions for preventing malaria mortality in children in plasmodium falciparum endemic areas. *International journal of epidemiology*, 39(suppl_1):i88–i101.

Ekvall, H. (2003). Malaria and anemia. *Current opinion in hematology*, 10(2):108–114.

Gülmezoglu, A. and Garner, P. (1998). Malaria in pregnancy in endemic areas (cochrane review). *The Cochrane Library*, (3).

Guyatt, H. L. and Snow, R. W. (2004). Impact of malaria during pregnancy on low birth weight in sub-saharan africa. *Clinical microbiology reviews*, 17(4):760–769.

Kamau, A., Nyaga, V., Bauni, E., Tsofa, B., Noor, A. M., Bejon, P., Scott, J. A. G., and Hammitt, L. L. (2017). Trends in bednet ownership and usage, and the effect of bednets on malaria hospitalization in the kilifi health and demographic surveillance system (khdss): 2008–2015. *BMC infectious diseases*, 17(1):720.

Kayentao, K., Garner, P., van Eijk, A. M., Naidoo, I., Roper, C., Mulokozi, A., MacArthur, J. R., Luntamo, M., Ashorn, P., Doumbo, O. K., et al. (2013). Intermittent preventive therapy for malaria during pregnancy using 2 vs 3 or more doses of sulfadoxine-pyrimethamine and risk of low birth weight in africa: systematic review and meta-analysis. *Jama*, 309(6):594–604.

KMIS (2015). Kenya malaria indicator survey 2015 [mis22] https://dhsprogram.com/pubs/pdf/mis22/mis22.pdf.

Kramer, M. S. (1987). Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the world health organization*, 65(5):663.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics-Theory and methods*, 26(6):1481–1496.

Macharia, P. M., Giorgi, E., Noor, A. M., Waqo, E., Kiptui, R., Okiro, E. A., and Snow, R. W. (2018). Spatio-temporal analysis of plasmodium falciparum prevalence to understand the past and chart the future of malaria control in kenya. *Malaria journal*, 17(1):340.

Menendez, C. (1995). Malaria during pregnancy: a priority area of malaria research and control. *Parasitology today*, 11(5):178–183.

Metgud, C. S., Naik, V. A., and Mallapur, M. D. (2012). Factors affecting birth weight of a newborn–a community based study in rural karnataka, india. *PloS one*, 7(7):e40040.

Mmbando, B. P., Cole-Lewis, H., Sembuche, S., Kamugisha, M., Theander, T., Lusingu, J., and Lemnge, M. (2008). Risk factors for low birth-weight in areas with varying malaria transmission in korogwe, tanzania: implications for malaria control. *Tanzania Journal of Health Research*, 10(3):137–143.

Mogeni, P., Williams, T. N., Fegan, G., Nyundo, C., Bauni, E., Mwai, K., Omedo, I., Njuguna, P., Newton, C. R., Osier, F., et al. (2016). Age, spatial, and temporal variations in hospital admissions with malaria in kilifi county, kenya: a 25-year longitudinal observational study. *PLoS medicine*, 13(6):e1002047.

Muchemi, O. M., Echoka, E., and Makokha, A. (2015). Factors associated with low birth weight among neonates born at olkalou district hospital, central region, kenya. *Pan African Medical Journal*, 20(1).

Mutabingwa, T. K., Bolla, M. C., Li, J.-L., Domingo, G. J., Li, X., Fried, M., and Duffy, P. E. (2005). Maternal malaria and gravidity interact to modify infant susceptibility to malaria. *PLoS Med*, 2(12):e407.

Nosten, F., McGready, R., and Mutabingwa, T. (2007). Case management of malaria in pregnancy. *The Lancet infectious diseases*, 7(2):118–125.

Nosten, F., Ter Kuile, F., Maelankirri, L., Decludt, B., and White, N. (1991). Malaria during pregnancy in an area of unstable endemicity. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 85(4):424–429.

Okiring, J., Olwoch, P., Kakuru, A., Okou, J., Ochokoru, H., Ochieng, T. A., Kajubi, R., Kamya, M. R., Dorsey, G., and Tusting, L. S. (2019). Household and maternal risk factors for malaria in pregnancy in a highly endemic area of uganda: a prospective cohort study. *Malaria journal*, 18(1):144.

Okiro, E. A., Hay, S. I., Gikandi, P. W., Sharif, S. K., Noor, A. M., Peshu, N., Marsh, K., and Snow, R. W. (2007). The decline in paediatric malaria admissions on the coast of kenya. *Malaria journal*, 6(1):151.

O'Meara, W. P., Bejon, P., Mwangi, T. W., Okiro, E. A., Peshu, N., Snow, R. W., Newton, C. R., and Marsh, K. (2008). Effect of a fall in malaria transmission on morbidity and mortality in kilifi, kenya. *The lancet*, 372(9649):1555–1562.

Ouma, P., Van Eijk, A. M., Hamel, M. J., Parise, M., Ayisi, J. G., Otieno, K., Kager, P. A., and Slutsker, L. (2007). Malaria and anaemia among pregnant women at first antenatal clinic visit in kisumu, western kenya. *Tropical Medicine & International Health*, 12(12):1515–1523.

Peters, P. J., Thigpen, M. C., Parise, M. E., and Newman, R. D. (2007). Safety and toxicity of sulfadoxine/pyrimethamine. *Drug safety*, 30(6):481–501.

Robert, B., Alec, I., and Thomson, P. (2004). The world health report: 2004: changing history.

Rogerson, S. J. (2017). Management of malaria in pregnancy. *The Indian journal of medical research*, 146(3):328.

Rogerson, S. J., Hviid, L., Duffy, P. E., Leke, R. F., and Taylor, D. W. (2007). Malaria in pregnancy: pathogenesis and immunity. *The Lancet infectious diseases*, 7(2):105–117.

Sen, P. K. (1968). Estimates of the regression coefficient based on kendall's tau. *Journal of the American statistical association*, 63(324):1379–1389.

Shulman, C., Marshall, T., Dorman, E., Bulmer, J., Cutts, F., Peshu, N., and Marsh, K. (2001). Malaria in pregnancy: adverse effects on haemoglobin levels and birthweight in primigravidae and multigravidae. *Tropical Medicine & International Health*, 6(10):770–778.

Snow, R., Peshu, N., Forster, D., Mwenesi, H., and Marsh, K. (1992). The role of shops in the treatment and prevention of childhood malaria on the coast of kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 86(3):237–239.

Snow, R. W., Kibuchi, E., Karuri, S. W., Sang, G., Gitonga, C. W., Mwandawiro, C., Bejon, P., and Noor, A. M. (2015). Changing malaria prevalence on the kenyan coast since 1974: climate, drugs and vector control. *Plos one*, 10(6):e0128792.

Steketee, R. W., Wirima, J. J., Hightower, A. W., Slutsker, L., Heymann, D. L., and Breman, J. G. (1996). The effect of malaria and malaria prevention in pregnancy on offspring birthweight, prematurity, and intrauterine growth retardation in rural malawi. *The American journal of tropical medicine and hygiene*, 55(1_Suppl):33–41.

Unger, H. W., Rosanas-Urgell, A., Robinson, L. J., Ome-Kaius, M., Jally, S., Umbers, A. J., Pomat, W., Mueller, I., Kattenberg, E., and Rogerson, S. J. (2019). Microscopic and submicroscopic plasmodium falciparum infection, maternal anaemia and adverse pregnancy outcomes in papua new guinea: a cohort study. *Malaria journal*, 18(1):302.

Walker, P. G., ter Kuile, F. O., Garske, T., Menendez, C., and Ghani, A. C. (2014). Estimated risk of placental infection and low birthweight attributable to plasmodium falciparum malaria in africa in 2010: a modelling study. *The Lancet Global Health*, 2(8):e460–e467.

Wardlaw, T. M. (2004). *Low birthweight: country, regional and global estimates*. Unicef.

Were, F., Mukhwana, B., and Musoke, R. (2002). Neonatal survival of infants less than 2000 grams born at kenyatta national hospital. *East African medical journal*, 79(2):77–79.

WHO (2019). World malaria report 2019. geneva: World health organization; 2019.