Geo-statistical modelling of malaria mortality and its relationship with anaemia in the Kisumu health and demographic surveillance system, Kenya

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät der
Universität Basel

von
Sammy Khagayi
Aus Tiriki, Kenya

Basel, 2024

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Jürg Utzinger (Fakultätsverantworlicher)
PD Dr. Penelope Vounatsou (Dissertationsleiter)
Dr, Pauline Mwinzi (Korreferent).

Basel, den 17/10/2023

Prof. Dr. Marcel Mayor
Dekan

Dedicated to the memory of my late Father, **KHAGGAYI Lawrence**!

Table of contents

List of abbreviations

ACT	Artemisinin-based Combination Therapy
AIDS	Acquired Immune Deficiency Syndrome
BCI	Bayesian Credible Interval
CDC	Centers for Disease Control and Prevention
CGHR	Centre for Global Health Research
CI	Credible Intervals
DDT	Dichlorodiphenyltrichloroethane
DHS	Demographic and Health Survey
DIC	Deviance Information Criteria
DNMP	Division of National Malaria Programme
HDSS	Health and Demographic Surveillance Systems
EIR	Entomological Inoculation Rate
EROS	Earth Resources Observation and Science
EVI	Enhanced Vegetation Index
ERC	European Research Council
GBD	Global Burden of Disease
GIS	Geographical Information System
HIV	Human Immunodeficiency Virus
Hb	Haemoglobin
HBR	Human Biting Rate
INDEPTH Network	International Network for the Demographic Evaluation of Populations and Their Health
IPT	Intermittent Preventive Treatment
IRB	Institutional Review Board
ITN	Insecticide Treated Net
KEMRI	Kenya Medical Research Institute
KHDSS	KEMRI Health and Demographic Surveillance Systems
LC	Land cover
LLIN	Long Lasting Insecticide-treated Nets
LMIC	Low- and Middle-Income Countries
LST	Land Surface Temperature
OR	Odds Ratio

MAP	Malaria Atlas Project
MCA	Multiple Correspondence Analysis
MCMC	Markov Chain Monte Carlo
MIS	Malaria Indicator Surveys
MITS	Minimally Invasive Tissue Sampling
MODIS	Moderate Resolution Imaging Spectroradiometer
MOH	Ministry of Health
MOPHS	Ministry of Public Health and Sanitation
MTIMBA	Malaria Transmission Intensity and Mortality Burden across Africa
PAF	Population Attributable Fraction
PhD	Doctor of Philosophy
PID	Personal Identification Number
PP	Parasite Prevalence
PRR	Prevalence Rate Ratios
PYO	Person Years of Observation
RBM	Roll Back Malaria Partnership
RDT	Rapid Diagnostic Test
RS	Remote Sensing
RR	Rate Ratios
SDG	Sustainable Development Goal
SES	Socio-Economic Status
SERU	Scientific and Ethics Review Unit
SIR	Sporozoite Infection Rate
SNF	Swiss National Foundation
SPR	Slide Positivity Rate
SSA	Sub-Saharan Africa
STH	Soil Transmitted Helminths
Swiss TPH	Swiss Tropical and Public Health Institute
TB	Tuberculosis
VA	Verbal Autopsy
USGS FEWS	United States Geological Surveys' Famine Early Warning Systems
WHO	World Health Organization

List of figures

List of tables

Summary

Malaria remains a significant global health concern, with sub-Saharan Africa bearing the brunt, especially among children under five. Despite efforts to reduce its prevalence, challenges persist, Kenya, particularly its western regions, sees over a third of its outpatient visits due to malaria. Although the disease has declined especially in the last two decades, progress has plateaued. Inconsistencies in data, lack of precise regional estimates, and inadequate monitoring are among the causes attributed to the stagnation. Effective data management and utilization are therefore vital in informing timely and effective policy decisions.

Our research, focused on a Health and Demographic Surveillance System (which is a population surveillance cohort), located in western Kenya, an area with high malaria prevalence, between the year 2007 and 2015. We investigated the spatio-temporal dynamics of malaria incidence and prevalence (as measures of transmission) and their effects on mortality while adjusting for climatic, environmental and other associated factors across all age categories. Furthermore, we investigated the association between malaria vis-à-vis parasitic worms' prevalence on anaemia risk across different age groups in order to understand the contribution of malaria to anaemia in this region.

In the chapter 2, we investigated the spatial and temporal patterns of malaria incidence among children below 5 years using Bayesian hierarchical negative binomial models. Between 2007 and 2012, there was a notable decline in malaria incidence, from 775 cases to 540 cases per 1,000 person years of observation (pyo). Monthly incidence exhibited significant seasonal variations, with peaks during or immediately after the long and short rainy seasons. Enhanced Vegetation Index (EVI), socio-economic status, altitude, and the study region were all critical in influencing malaria risk. Meanwhile, climatic factors i.e. temperature and rainfall known to be critical in malaria transmission showcased non-significant associations in this study.

In the chapter 3, we explored the relationship between malaria incidence, through slide positivity rates (SPR) captured at health facilities around the KHDSS and mortality. We fit Bayesian spatio-temporal survival models to investigate the relation between mortality (all-cause/malaria-specific) and malaria incidence across all age groups. We found that a small increase in SPR was a driver of both all-cause/malaria-specific mortality. This was especially elevated in children aged 1-4 years. Interestingly, for older children aged 5-14 years, the study found a reduced association between SPR and malaria-specific mortality.

In the fourth chapter, using parasite prevalence data collected at the households as the metric for malaria transmission, we extended the investigation to its impact on mortality. Data from this study revealed that malaria parasitaemia was associated with mortality across age groups over four-to-five-year periods. While clinical malaria showcased a strong correlation with mortality across all age groups, most notably in children aged 5-14 years, the effect was not observed in neonates, adults, and the elderly.

The fifth chapter shifted to investigate anaemia prevalence and its link with malaria or parasitic worm infestations. We found a high and consistent prevalence of anaemia across all ages, predominantly impacting the youngest (1-11 months) and the elderly (65+). Meanwhile, malaria parasitaemia alone and clinical malaria were linked with a higher anaemia risk. There was no evidence of any effect of Schistosomiasis and helminths risk with anaemia prevalence.

All these studies converge on the need for an understanding of the multifaceted nature of malaria and related conditions in Western Kenya. While the initial focus is to map out the incidence of malaria and its ties with environmental and socio-economic factors, the subsequent studies amplify the importance of assessing the broader implications of malaria on mortality and co-morbidities like anaemia. The consistent use of the Health and Demographic Surveillance System (HDSS) data across studies ensures a coherent analysis.

From these results, we propose that SPR is a more efficient measure of gauging malaria transmission and its potential impact on mortality in an area of high endemicity. In essence,

while the global fight against malaria has made significant strides, collectively, these findings stress the need for targeted, multi-pronged intervention strategies to combat the complexities and repercussions of malaria in the region. By recognizing and addressing the unique challenges posed by specific regions, localized pockets like the one studied in western Kenya require urgent, specialized attention for more effective policy formulation and implementation.

Zusammenfassung

Malaria ist nach wie vor ein wichtiges globales Gesundheitsproblem, wobei die afrikanischen Länder südlich der Sahara die Hauptlast tragen, insbesondere bei Kindern unter fünf Jahren. Trotz der Bemühungen, die Prävalenz der Krankheit zu verringern, gibt es nach wie vor Probleme. In Kenia, insbesondere in den westlichen Regionen, wird mehr als ein Drittel der ambulanten Behandlungen auf Malaria zurückgeführt. Obwohl die Krankheit vor allem in den letzten beiden Jahrzehnten zurückgegangen ist, sind die Fortschritte ins Stocken geraten. Unstimmigkeiten bei den Daten, das Fehlen genauer regionaler Schätzungen und eine unzureichende Überwachung gehören zu den Ursachen, die für die Stagnation verantwortlich gemacht werden. Eine wirksame Datenverwaltung und -nutzung ist daher von entscheidender Bedeutung, um zeitnahe und wirksame politische Entscheidungen treffen zu können.

Unsere Forschung konzentrierte sich auf ein Health and Demographic Surveillance System (eine Bevölkerungsüberwachungskohorte) im Westen Kenias, einem Gebiet mit hoher Malariaprävalenz, zwischen 2007 und 2015. Wir untersuchten die räumlich-zeitliche Dynamik der Malariainzidenz und -prävalenz (als Maß für die Übertragung) und ihre Auswirkungen auf die Sterblichkeit, wobei wir klimatische, umweltbedingte und andere damit verbundene Faktoren in allen Alterskategorien berücksichtigten. Darüber hinaus untersuchten wir den Zusammenhang zwischen der Malariaprävalenz und der Prävalenz parasitärer Würmer und dem Anämierisiko in verschiedenen Altersgruppen, um den Beitrag der Malaria zur Anämie in dieser Region zu verstehen.

In Kapitel 2 untersuchten wir die räumlichen und zeitlichen Muster des Auftretens von Malaria bei Kindern unter 5 Jahren mit Hilfe von hierarchischen negativen Binomialmodellen nach Bayes. Zwischen 2007 und 2012 gab es einen bemerkenswerten Rückgang der Malariainzidenz von 775 Fällen auf 540 Fälle pro 1.000 Personenjahren Beobachtung (pyo). Die monatliche

Inzidenz wies erhebliche saisonale Schwankungen auf, mit Spitzenwerten während oder unmittelbar nach den langen und kurzen Regenzeiten. Der Enhanced Vegetation Index (EVI), der sozioökonomische Status, die Höhenlage und die Untersuchungsregion spielten eine entscheidende Rolle für das Malariarisiko. Klimafaktoren, d. h. Temperatur und Niederschlag, die bekanntermaßen für die Malariaübertragung entscheidend sind, wiesen in dieser Studie keine signifikanten Zusammenhänge auf.

In Kapitel 3 untersuchten wir die Beziehung zwischen der Malaria-Inzidenz, die anhand der in den Gesundheitseinrichtungen rund um das KHDSS erfassten Slide-Positivity-Raten (SPR) ermittelt wurde, und der Mortalität. Mit Hilfe von Bayes'schen räumlich-zeitlichen Überlebensmodellen untersuchten wir die Beziehung zwischen der Sterblichkeit (alle Ursachen/Malaria-spezifisch) und der Malaria-Inzidenz in allen Altersgruppen. Wir fanden heraus, dass ein geringer Anstieg der SPR sowohl die Gesamtmortalität als auch die malariaspezifische Mortalität beeinflusst. Besonders ausgeprägt war dies bei Kindern im Alter von 1-4 Jahren. Interessanterweise ergab die Studie für ältere Kinder im Alter von 5-14 Jahren einen geringeren Zusammenhang zwischen SPR und malariaspezifischer Sterblichkeit.

Im vierten Kapitel wurden die in den Haushalten erhobenen Daten zur Parasitenprävalenz als Maßstab für die Malariaübertragung herangezogen und die Untersuchung auf ihre Auswirkungen auf die Sterblichkeit ausgeweitet. Die Daten dieser Studie zeigten, dass die Malariaparasitämie in allen Altersgruppen über einen Zeitraum von vier bis fünf Jahren mit der Sterblichkeit in Zusammenhang stand. Während die klinische Malaria in allen Altersgruppen eine starke Korrelation mit der Sterblichkeit aufwies, insbesondere bei Kindern im Alter von 5-14 Jahren, wurde dieser Effekt bei Neugeborenen, Erwachsenen und älteren Menschen nicht beobachtet.

Im fünften Kapitel untersuchten wir die Prävalenz von Anämie und deren Zusammenhang mit Malaria oder parasitärem Wurmbefall. Wir fanden eine hohe und konsistente Prävalenz von Anämie in allen Altersgruppen, wobei vor allem die Jüngsten (1-11 Monate) und die älteren

Menschen (65+) betroffen waren. Gleichzeitig waren die Malariaparasitämie allein und die klinische Malaria mit einem höheren Anämierisiko verbunden. Es gab keine Hinweise auf eine Auswirkung des Schistosomiasis- und Helminthenrisikos auf die Anämieprävalenz.

Alle diese Studien zeigen, dass ein Verständnis für die Vielschichtigkeit der Malaria und der damit zusammenhängenden Erkrankungen im Westen Kenias erforderlich ist. Während der anfängliche Schwerpunkt darauf liegt, die Häufigkeit von Malaria und ihre Verbindungen zu umweltbedingten und sozioökonomischen Faktoren zu erfassen, unterstreichen die nachfolgenden Studien die Bedeutung der Bewertung der umfassenderen Auswirkungen von Malaria auf die Mortalität und Komorbiditäten wie Anämie. Die einheitliche Verwendung der Daten des Health and Demographic Surveillance System (HDSS) in allen Studien gewährleistet eine kohärente Analyse.

Auf der Grundlage dieser Ergebnisse schlagen wir vor, dass die SPR ein effizienteres Maß für die Messung der Malariaübertragung und ihrer potenziellen Auswirkungen auf die Sterblichkeit in einem Gebiet mit hoher Endemie darstellt. Auch wenn der weltweite Kampf gegen Malaria erhebliche Fortschritte gemacht hat, unterstreichen diese Ergebnisse insgesamt die Notwendigkeit gezielter, mehrgleisiger Interventionsstrategien, um die Komplexität und die Auswirkungen von Malaria in der Region zu bekämpfen. Durch die Anerkennung und Bewältigung der einzigartigen Herausforderungen, die sich in bestimmten Regionen stellen, erfordern lokal begrenzte Gebiete wie das untersuchte in West-Kenia dringende, spezielle Aufmerksamkeit für eine effektivere Politikformulierung und -umsetzung.

Acknowledgements

I wish to acknowledge a number of people who have contributed to this thesis either directly or indirectly. Sincerely grateful to PD. Dr. Penelope Vounatsou for the mentorship and guidance in the field of Bayesian statistics and geo-spatial modelling. Prof. Dr Juerg Utzinger, director Swiss TPH has had immense help in finalising this journey. Special thanks to Dr Pauline Mwinzi, the co-referee for this thesis.

Many thanks for the immense support from Christine Mensch and the EPH secretariat who through their assistance, made the stay in Basel so easy.

To Astrid thanks for everything! I may never be able to repay you back! Special mention dear friends and colleagues, Diboulo, Dlamini, Nambuusi, Ssempiira, Massoda-Tonye, Millogo, Belocconi, Nyawanda, Mhimbira, Ntuku, Koku-Awonoor, Kokaliaris, Bärenbold, and Malinga. Not to forget the many special people not mentioned in name, for the camaraderie in and around Basel, you are highly appreciated.

Special thanks to Kwaro, Amek and others back in Kenya who have really stuck out their necks for me. I am indebted to the KEMRI-CGHR, for allowing me to use the platform and be away for extended periods of time!

To my mum, Ann, Christine, Ronnie, Claire and your families' thanks for moral support, prayers and encouragement.

My wife **Sally** and daughters **Maria** and **Carrie**, it has been tough but we have more precious times ahead.

This PhD was supported by the European Research Council (ERC) fellowship of Penelope Vounatsou grant number 323180 and the Swiss National Foundation (SNF) program for Research on Global Issues for Development (R4D) project number IZ01Z0.

1. Chapter 1: Introduction and objectives

1.1. Introduction and background

Malaria remains a leading cause of morbidity and mortality in the world with an estimated 3.2 billion people at risk. The brunt of this burden falls heavily on sub-Saharan Africa (SSA) which accounts for over 95% of malaria cases and 96% of related death (Weiss et al., 2019; World Health Organization, 2022) (Figure 1.1). In this region, children under five years make up 76% of the deaths due to malaria (World Health Organization, 2022). Such statistics have hindered the achievement of the goal to reduce malaria morbidity and mortality by 75% by the year 2025 (World Health Organization, 2021). However, malaria is preventable and treatable if the correct combination of prevention, treatment and surveillance strategies are put in place (Cibulskis et al. 2011). For effective policymaking, especially in low-income SSA countries, there is a need for accurate tracking of progress in malaria control through the establishment of efficient surveillance systems that will ensure reliable data collection, analysis and dissemination.

In Kenya, there has been a considerable reduction in malaria morbidity and mortality since the late 20th century (Division of National Malaria Programme (DNMP) & ICF, 2021; Noor, Gething, et al., 2009). However, malaria still accounts for over one-third of all outpatient visits. The risk of infection is highest in some parts of western country with an estimated parasitaemia prevalence of 38% around the Lake Victoria endemic region (Division of National Malaria Programme (DNMP) & ICF, 2021; Khagayi et al., 2019).

Despite substantial investment in malaria control initiatives both domestically and internationally (World Health Organization, 2022), along with global advances in prevention, diagnostics and treatment (Bhatt et al., 2015; Hanboonkunupakarn & White, 2022), the rate of progress in reducing malaria has plateaued. Key challenges include inconsistent data, imprecise estimates at the sub-regional level, and inadequate surveillance systems to track localized outbreaks (Dhiman, 2019). Additionally, even when data exists, it often remains underutilized

due to inadequate management and analysis (Okello et al., 2019). It is therefore critical to optimally leverage all available data resources, analyse it and disseminate results in order to best guide policy decisions locally, regionally or globally in a timely manner.

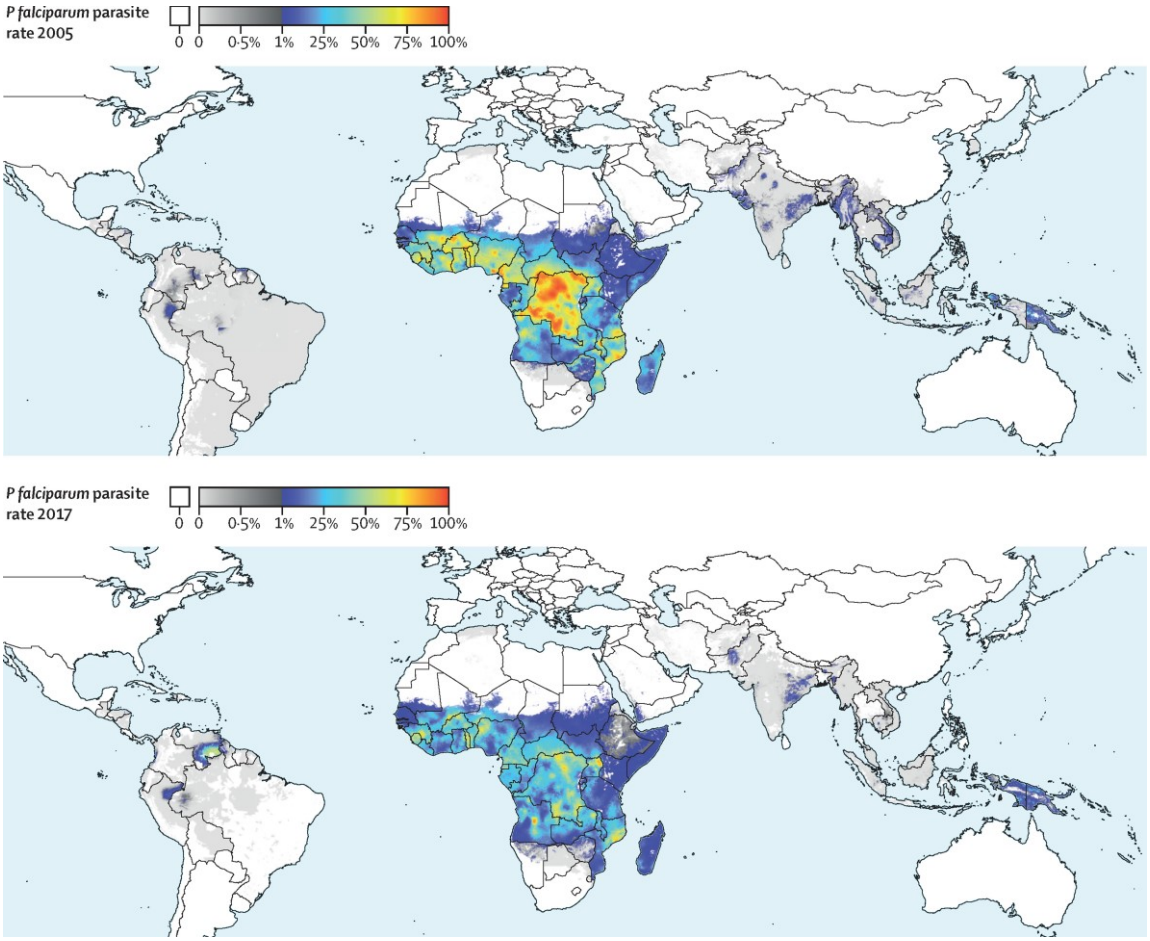


Figure 1.1: Trends in global *Plasmodium falciparum* endemicity (Adapted from, Weiss et al 2019)

1.2. The life cycle and transmission of malaria

Malaria is a vector borne disease caused by protozoan parasites of the genus *plasmodium.* There are four main plasmodia species known to cause malaria in humans; *Plasmodium falciparum, P. vivax, P. malariae* and *P. ovale*, with *P. falciparum* being the most common one in SSA and responsible for most deaths and morbidity (Greenwood et al. 2005; Snow and Omumbo 2006). However, in recent years cases of malaria in humans have been reported due to a fifth species *P. knowlesi* which is previously known to cause malaria among primates in certain areas of South-East Asia (Amir et al., 2018).

The malaria parasite life cycle is complex and involves multiple stages, both within humans and mosquitoes (Figure 1.2). The malaria parasite is transmitted to a human being through a bite from an infected mosquito. The mosquito transfers sporozoites (last stage of the parasite in the mosquito) during feeding into the human blood stream. These sporozoites then move to the liver, where they multiply asexually to become merozoites. These merozoites are released into the bloodstream where they either develop into male and female gametocytes, or, under repeated replication, form more merozoites which result in symptomatic disease. A subsequent mosquito bite on an infected human ingest these gametocytes (male/female). In the mosquito, the gametocytes develop into gametes that sexually reproduce to form zygotes. The zygotes mature into oocyst and burst, which releases sporozoites that eventually get into the mosquitos' salivary glands. If this infected mosquito bites another human, they are transferred into another human and another life cycle starts (Beier, 1998; Greenwood et al., 2005; White et al., 2014).

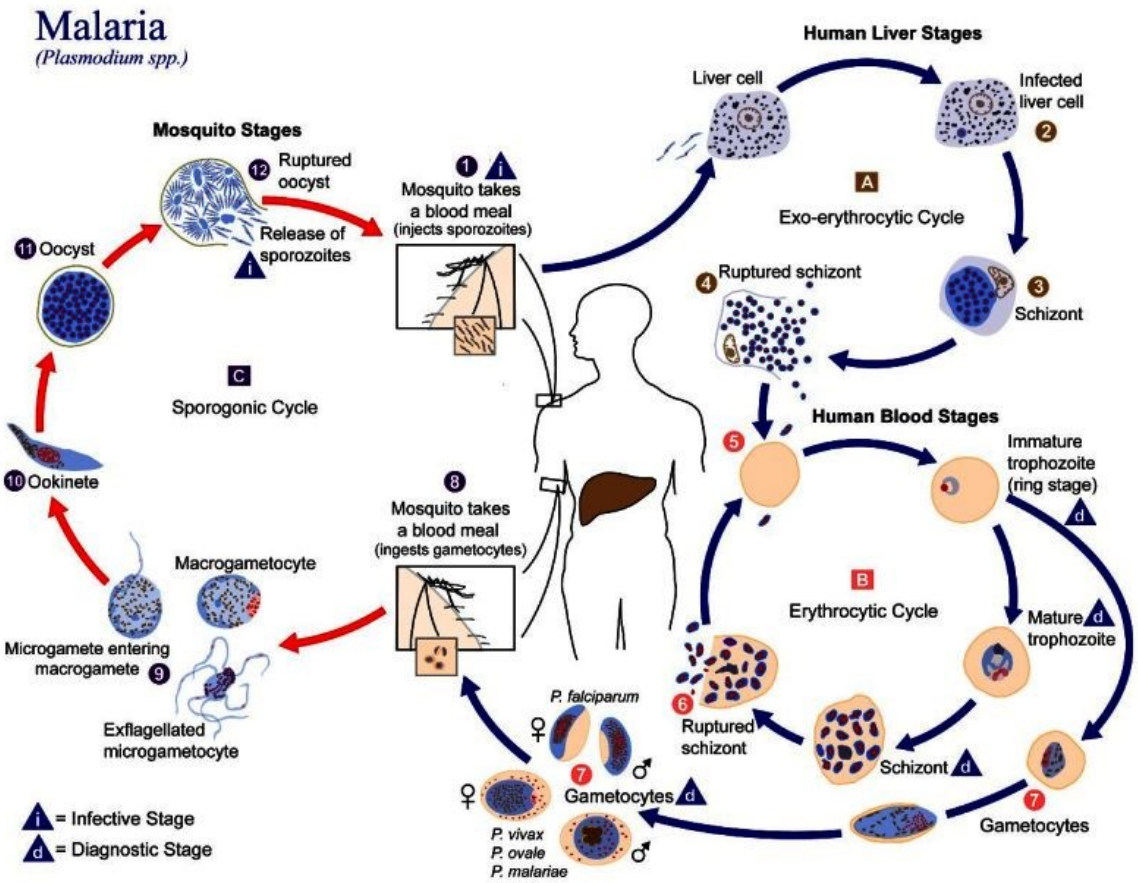


Figure 1.2: Life cycle of the malaria parasite (Adapted from the CDC-Global Health, Division of Parasitic Diseases and Malaria)

The parasite`s completion of its life cycle and ultimately determination of transmission and morbidity is a complex process driven largely by various environmental, topographical and climatic factors (Githeko et al., 2006; Zhou et al., 2004).

Measuring malaria transmission is crucial for understanding its overall impact. Various metrics have been used to do this; including human infection rates such as parasite prevalence, incidence rates, and prevalence rates. Additionally, the interaction between mosquitoes and humans as measured using the entomological inoculation rate (EIR) defined as the product of the human biting rate (HBR) and sporozoite infection rate (SIR) or vector measures (mosquito density) have also been used.

1.3. **Malaria mortality burden**

Malaria parasite infection can lead to an array of health complications ranging from clinical manifestation of fever, convulsions, anaemia, respiratory distress, cerebral malaria, organ failures, coagulation disorders, electrolyte imbalances or splenomegaly (Weatherall et al., 2002). Additionally, malaria often coincides with other medical conditions and infections such as bacteremia, sepsis, pregnancy complications and malnutrition (Bartoloni & Zammarchi, 2012; World Health, 2000). Without timely and effective treatment, malaria can easily escalate to severe illness and ultimately, death (Bartoloni & Zammarchi, 2012; Miller et al., 2002; Phillips et al., 2017).

Beyond its impact on individual health, malaria has been shown to negatively affect the social and economic wellbeing of families and society. With the disease even linked to poverty perpetuation (Sachs & Malaney, 2002). These complex manifestations and implications make it challenging to estimate the complete burden of malaria, the extent of malaria's toll, particularly in terms of its contribution to mortality rates, especially in sub-Saharan Africa and

notably in Kenya. Numerous studies have highlighted the significant burden of malaria-related mortality in these regions (Murray et al., 2012; Rumisha et al., 2014; World Health Organization, 2021). In Kenya, malaria remains the leading cause of death, most prominently among children under the age of five (Division of National Malaria Programme (DNMP) & ICF, 2021).

1.3.1. Evaluating the relationship between malaria transmission measures and mortality

Estimation of malaria cases and deaths globally is carried out using different methods. In estimating malaria-mortality, given its complex nature, cases and deaths are used separately or in combination with entomological measures to come up with either incidence, prevalence or EIR to measure trends (Amek et al., 2018; Murray et al., 2012; Weiss et al., 2019; WHO Malaria Policy Advisory Committee and Secretariat, 2012). However, these studies and reports use different methods, data and assumptions in doing so.

The World Health Organization (WHO) employs a multipronged strategy for estimating malaria cases, contingent on factors like the endemic nature of the region, quality of surveillance systems, and data availability (World Health Organization, 2022). In most SSA countries with poor reporting systems, malaria incidence is predicted using models based on climatic predictors and adjusted over space and time (Alegana et al., 2016; Cibulskis et al., 2011; Gething et al., 2016; Weiss et al., 2019). Among the recommendations by the WHO's working group is that future estimates of malaria burden should include or adjust for insecticide treated mosquito nets (ITNs) and long-lasting insecticide-treated nets' (LLINs) effectiveness. There should also be consideration of slide positivity among febrile children whether seeking care or not. And lastly, future estimations should also take into consideration the time series prevalence of *P. falciparum* in relation to seasonality (WHO Malaria Policy Advisory Committee and Secretariat, 2013).

Malaria diagnosis is widely done based on presence of parasites using rapid diagnostic tests or microscopy in persons with reported fever (Zimmerman & Howes, 2015). These are widely available and cost-effective especially in resource limited settings, where the disease is most experienced. With proper diagnosis, malaria incidence and prevalence can thus be established. However, some people would have the malaria parasite but do not exhibit any signs and symptoms of illness (Bousema et al., 2014; Laishram et al., 2012); therefore, disentangling the contribution of malaria parasites presence alone or clinical malaria to mortality is paramount.

Cause of death identification due to malaria has also been quite challenging. Apart from the difficulty in diagnosing malaria as the cause of illness, many deaths occur at home in many low-income settings make it complex to confirm malaria as a cause of death (Adair, 2021; Hamel et al., 2011; Leitao et al., 2013). Verbal autopsies (VA), developed through the years, have been useful in determining cause of death using different tools and methodologies (Byass et al., 2019; Nichols et al., 2018). Despite some inherent shortcomings, especially the sensitivity of diagnosing malaria using VA (Herrera et al., 2017; Rakislova et al., 2021), there have been improvements over the years. Making it the best available tool for cause of death determination at population level (Bailo et al., 2022; Thomas et al., 2018) from which malaria-related deaths can be extracted for burden estimation.

Another important aspect to consider when estimating malaria-related mortality is that climatic and environmental factors are huge drivers of the parasite development life cycle and hence malaria infection. Nowadays, climatic and environmental proxies at high spatial resolution are available from remote sensing (RS) and have been used to predict disease burden estimates at locations without data (Diboulo et al., 2015; Ssempiira et al., 2018; Weiss et al., 2019).

In Kenya, malaria prevalence estimates have been obtained from malaria indicator surveys (MIS) carried out by the national malaria control program. Over the last two decades, results from these surveys indicate a reduction in malaria attributed morbidity, especially in children below the age of 5 years. The highest risk of infection is still experienced around the lake

endemic region in western Kenya (Division of National Malaria Programme (DNMP) & ICF, 2021). However, these surveys mostly rely on one-off surveys on a national sample. The data are also aggregated at national and subnational level while overlooking local spatial-temporal disparities, at the same time their estimates are not linked to mortality.

Bayesian geostatistical models that were used to estimate malaria risk in Kenya in 2009 observed similar trends to the MIS and identified the lake endemic region as having the biggest contribution to malaria in the country (Noor, Gething, et al., 2009). However, the main shortcoming of this study was the use of data from different sources which were not comparable and covered only one year.

In line with the WHO recommendations on malaria burden estimation, modelling at fine spatial resolution has in recent times explored the relationship between transmission and malaria specific mortality/morbidity (Amek et al., 2012; Kasasa et al., 2013; Rumisha et al., 2014). These studies took into consideration data well-aligned in space and time, used verbal autopsy for cause of death and considered climatic/environmental factors using longitudinal population follow-ups. While they emphasized the importance of localized data analysis, the use of spatial-temporal techniques and use of mortality data, they failed to account for the impact of various interventions on malaria incidence, prevalence and mortality.

1.3.2. Burden of malaria-related anaemia

Anaemia, a condition characterized by reduction of haemoglobin (Hb) levels in the blood has been linked to malaria in different settings (Menendez et al., 2000; Sankaran & Weiss, 2015; White, 2018). While the risk of anaemia varies from place to place, studies have shown that in areas of high malaria transmission severe anaemia is the most common presentation of *P. falciparum* infection and results in adverse outcomes (Menendez et al., 2000; Obonyo et al., 2007). The condition is especially widespread in low-income settings in SSA, where risk levels range from 47% to 67% (World Health Organization, 2015a).

In Kenya, the national malaria indicator survey (MIS) shows that prevalence of any anaemia in children under 5 years is approximately 52% nationally. This figures jump to 62% in the western Kenya region around Lake Victoria (Division of National Malaria Programme (DNMP) & ICF, 2021). In some age-specific studies in this region, it has even been suggested that malaria and anaemia should be combined into one group of disease as sometimes there is no clear disentanglement on the direction of causality (Desai et al., 2014; Hamel et al., 2011; Obonyo et al., 2007; Sewe et al., 2015). Furthermore, national trends in anaemia prevalence often mirror those of malaria (Division of National Malaria Programme (DNMP) & ICF, 2021). However, it is crucial to note that anaemia is influenced by other factors like helminths infection, schistosomiasis, malnutrition, infectious diseases and HIV/AIDS. Anaemia can also be due to genetic disorders like sickle cell disease, α-thalassaemia or G6PD deficiency among others. These are diverse, could coexist or may indicate the ability to resist malaria illness (Foote et al., 2013; GBD Anaemia Collaborators, 2023; Kassebaum, 2016; Soares Magalhães & Clements, 2011; World Health Organization, 2015a). Given these multiple contributing factors, understanding the specific role that malaria plays in causing anaemia in this region is essential in the elimination of malaria burden and prevention of anaemia.

1.4. **Health and demographic surveillance systems**

Many countries in Sub-Saharan Africa (SSA) face significant limitations in health data collection and use due to inadequate infrastructure and limited healthcare access (Byass, 2007; Fottrell, 2009; Koumamba et al., 2021). To address this, health and demographic surveillance systems (HDSS) sites were established in various developing countries. These systems were aimed at offering valuable insights into the demographic and epidemiological landscape of specific populations, albeit on a smaller scale (Bos, 2004; Sankoh & Byass, 2012). Over the years, these sites have expanded and included a lot more specialized and diverse data in their

routine activities. Making them an invaluable resource for health research in the developing world.

1.4.1. The KEMRI-CGHR health and demographic surveillance system

The Kenya Medical Research Institute's Centre for Global Health Research launched an HDSS site (KHDSS) in 2001. Established in partnership with the Centers for Disease Control and Prevention (CDC) and others, this site originally evolved from a randomized control trial focusing on bed net usage in Asembo, located in the current Siaya County of western Kenya. The KHDSS later expanded to include Gem and Karemo areas in the same county (Figure 1.3). As of mid-2015, the KHDSS monitored approximately 250,000 individuals through a continuous tri-annual surveys in this demographically-defined area (Adazu et al., 2005; Odhiambo et al., 2012).

Figure 1.3: Maps of the KEMRI-CGHR health and demographic surveillance system in western Kenya

Once an individual is included in the KHDSS through a census, in-migration, or birth, they are assigned a unique Personal Identification Number (PID) based on their village, household, and compound numbers. This PID remains constant throughout the study. The KHDSS collects and

updates a variety of data, such as birth and death records, migration patterns, and socioeconomic indicators.

The KHDSS site was involved in the Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) project that sought to analyse data on malaria transmission intensity and burden in Africa from 2002 to 2004. Since then, the site has continued to collect data on malaria indicators. The extensive dataset, covering clinical symptoms, confirmed malaria cases, and entomological factors, is an invaluable resource for estimating local disease burdens.

While HDSS sites like KHDSS offer crucial data, they have inherent limitations. First, they only cover a limited geographical area and thus may not fully represent broader populations. Second, the requirement for an individual to reside in the surveillance area for at least three months may skew the data towards more stable, less mobile residents. Third, accurately identifying and tracking individuals can pose challenges. Lastly, HDSS sites often facilitate additional health studies and interventions, potentially leading to better healthcare access for monitored populations, which may bias the data. Despite these drawbacks, the detailed and long-term data collected by HDSS sites are invaluable for research in resource-limited settings (Bos, 2004).

1.5. **Bayesian spatial-temporal modelling**

The KHDSS collects geo-referenced longitudinal data at household level, and due to this, the data is often highly correlated in space and time (Amek et al., 2011). Such data should therefore not be analysed using standard statistical methods that assume independence of observations, since such methods would cause the standard errors of estimated parameters to be over or underestimated (Cressie, 1993; Giardina et al., 2012; Gosoniu et al., 2012). Data from HDSS sites, collected over a long period of time and remote sensing data, contain large number of parameters that make classical analysis infeasible.

To address these challenges, Bayesian methods are often employed, particularly those that use Markov Chain Monte Carlo (MCMC) simulation algorithms (Gelfand & Smith, 1990; Robert & Casella, 2011). Advances in Bayesian computation have enabled integration of stochastic mathematical models into general statistical frameworks. This is particularly useful for spatial prediction and forecasting (Clancy & O'Neill, 2008; O'Hara & Sillanpää, 2009). These advancements have paved the way for large-scale statistical analyses in infectious disease studies, thereby informing public health decisions (Gemperli et al., 2006; Gething et al., 2016; Giardina et al., 2012).

In Bayesian modelling, available data is combined with prior beliefs about model parameters through likelihood functions. This approach allows for making informed inferences about the model's parameters based on existing evidence (known as posteriors). Bayesian spatial-temporal methods have been specifically applied to model malaria transmission and disease burden in various settings (Amek et al., 2011; Giardina et al., 2012; Kasasa et al., 2013; Rumisha et al., 2014; Ssempiira et al., 2017). These models yield high-resolution estimates of malaria risk and transmission dynamics, providing invaluable insights for public health interventions.

1.6. **Rationale of the research**

Getting precise mortality data is a long-standing issue in many developing countries, especially in pinpointing specific causes of death (Bos, 2004). This data limitation hampers evaluation of how effective malaria control interventions perform.

Earlier research within the KHDSS from 2002 to 2004 used Bayesian methods to investigate the relation EIR and mortality rates (Amek, 2013). The findings indicated that areas with higher malaria transmission also experienced elevated mortality rates, particularly among young children. Furthermore, they underscored the critical role that both time and space play in

understanding malaria dynamics. However, these earlier efforts were limited in their scope, covering only short time periods and omitting key variables like confirmed malaria cases and use of insecticide-treated bed nets (ITNs) among other interventions.

To address these gaps, this thesis builds upon that prior work to enrich existing knowledge and inform policy. This was done by;

1. Incorporating a broader range of malaria indicators as recommended by the World Health Organization's expert committee.
2. Expanding the dataset to span an additional seven years and thus allowing for a more robust analysis over a longer timeframe.
3. Examining a more comprehensive set of health outcomes related to malaria, such as rates of hospitalization, asymptomatic infections, and instances of anaemia.

To achieve these aims, we employed advanced Bayesian geostatistical spatio-temporal models that leverage the extended data set collected by the KHDSS. This approach will enable a more nuanced and comprehensive understanding of the malaria landscape, thereby informing more effective interventions and policy decisions.

1.7. **Objectives**

The primary aim of this research was to refine malaria burden estimates in the KHDSS by utilizing advanced modelling techniques for the years 2007 to 2015. The specific objectives were as follows:

1. To assess dynamics in malaria incidence among children under 5 in relation to space and time in the KHDSS.
2. To explore the association between malaria transmission-measured by incidence rates and overall as well as malaria-specific mortality rates in the KHDSS for all age categories.

3. To assess the spatio-temporal association between prevalence as a measure of transmission and malaria mortality in the KHDSS across all age groups.
4. To estimate, the anaemia related burden, due to malaria in the KHDSS across all age groups.

2. Chapter 2: Spatio-temporal effects of climatic and environmental factors on malaria incidence among children under 5 years in Western Kenya.

Sammy Khagayi[1,2,3], Brian Nyawanda[1,2,3], Simon Kariuki[1], Stephen Munga[1], and Penelope Vounatsou[2,3§].

1. Kenya Medical Research Institute - Center for Global Health Research, Kisumu, Kenya
2. Swiss Tropical and Public Health Institute, Basel, Switzerland
3. University of Basel, Switzerland

[§]Corresponding author: sammy.khagayi@unibas.ch

This paper has been prepared for Plos Global Health

Abstract

Background
Malaria remains a significant public health concern in Western Kenya, particularly among children under 5 years old. A detailed understanding of the disease's spatio-temporal patterns and its associated risk factors is critical for targeted intervention in this vulnerable group.

Methods
Utilizing Bayesian hierarchical negative binomial models, we assessed the effect of climatic and environmental variables, including temperature, rainfall, and Enhanced Vegetation Index (EVI), among other factors on malaria incidence in children under 5 years within a Health and Demographic Surveillance System (HDSS) site located in western Kenya between 2007 and 2012.

Results
Overall, malaria incidence rates averaged 628 cases per 1,000 person years of observation (pyo), peaking at 775 cases in 2009 and declining to 540 cases by 2012. Monthly incidence exhibited significant seasonal fluctuations, particularly during May to July and November to January. Enhanced Vegetation Index (EVI) closely aligned with incidence peaks and showed a strong association, suggesting a composite effect of optimal climatic conditions for malaria transmission. An increase in mean temperature was associated with a decline in incidence while rainfall was associated with increased incidence, however they were not statistically important in their association. Socio-economic status, study area, and altitude also emerged as statistically important factors.

Conclusion
The study underscores the need for localized, timely preventive and control measures given the higher incidence rates and observed spatio-temporal variations. Socio-economic and environmental factors significantly influence the risk of malaria, emphasizing the necessity for a multi-pronged approach to malaria control**.**

Keywords: Malaria mortality, Incidence, Bayesian spatio-temporal, climate, environment Health and demographic surveillance system, Western Kenya

2.1. Introduction

There have been substantial strides globally in reduction of malaria-associated mortality and morbidity. However, malaria remains a significant contributor to the global disease burden, particularly in sub-Saharan Africa (SSA) (Bhatt et al., 2015; Gething et al., 2016; World Health Organization, 2022). In Kenya, the disease accounts for about 15% of all outpatient visits (Division of National Malaria Programme (DNMP) & ICF, 2021). In western Kenya, areas surrounding Lake Victoria have the highest burden in the country with a parasiteamia prevalence of 36% (Khagayi et al., 2019). While there has been notable progress in decreasing the burden of malaria in the hardest-hit countries, there are still difficulties in precisely quantifying and reporting the disease's prevalence, incidence, and annual trends (Gething et al., 2016; World Health Organization, 2022). These challenges hinder effective planning and slow down advancements in combating malaria.

The World Health Organization (WHO) advocates for surveillance as a key strategy to reduce the global impact of malaria by 2030, particularly in regions with high transmission rates. This is most crucial for children under the age of five, who are disproportionately affected (World Health Organization, 2021). With more than 70% of the population at risk for malaria in Kenya, children under five are also the most severely impacted (Division of National Malaria Programme (DNMP) & ICF, 2021).

For accurate assessment of malaria infection and its impact, it is crucial to have timely and precise data on when and where cases occur. However, this has often been hampered by ineffective surveillance systems that yield unreliable data, typically based only on confirmed cases from hospitals (Alegana et al., 2020; Gething et al., 2006). To capture the extent of malaria's burden well, it is essential to consider both the spatial and temporal dimensions of its incidence. This requires rigorous attention to detail, using the appropriate denominator, and accounting for healthcare utilization as well as under-reporting of cases (Cibulskis et al., 2011).

Additionally, malaria's relationship with climate and environmental factors adds another layer of complexity to its measurement (Amek et al., 2012). Fortunately, advancements in remote sensing technologies now provide easier access to these variables. By incorporating climate and environmental factors along with other relevant covariates in models, we can produce estimates that are more accurate. This, in turn, supports the development of more effective, localized policies for malaria control and eradication (Alegana et al., 2016; Beloconi et al., 2023; Nyawanda et al., 2023).

In this context, Health and Demographic Surveillance System (HDSS) sites provide the often-missing denominator that is crucial for accurate estimates, particularly in many parts of SSA, where functional vital registration systems are lacking (Sankoh & Byass, 2012). When paired with a well-organized system for tracking illness and relevant climatic data, HDSS sites can become invaluable tools for gauging the burden of diseases like malaria, especially in settings with limited resources.

We applied Bayesian hierarchical negative binomial models to explore how climate and environmental factors influence the incidence of malaria in children under five. This study was conducted at a long-established HDSS site in western Kenya from 2007 to 2012. By applying advanced statistical methodology to monthly data at the lowest administrative level, we offer a robust analytical framework for understanding and addressing a disease that continues to have a significant impact on global health.

2.2. Materials and methods

2.2.1. KEMRI-CGHR HDSS Profile

The Kenya Medical Research Institute-Center for Global Health Research (KEMRI-CGHR) in collaboration with other partners has been conducting in-depth studies and monitoring diseases in western Kenya for over thirty years. In 2001, they established a health and demographic

surveillance system (KHDSS) (Figure 2.1) in the Asembo area of Siaya County, western Kenya, which expanded to Gem the following year (2002) and later on to Karemo in 2007 (Adazu et al., 2005; Hamel et al., 2011; Odhiambo et al., 2012).

Located on the northern shores of Lake Victoria, the KHDSS monitors a population of more than 240,000 individuals across 58,700 households. It is in a malaria endemic zone with high prevalence of malaria and HIV (Khagayi et al., 2019; National AIDS and STI Control Programme (NASCOP), 2020). Within the KHDSS area, malaria is the primary cause of death and hospital admissions for children under 5 years old (Amek et al., 2014; Desai et al., 2014; Kwambai et al., 2023).

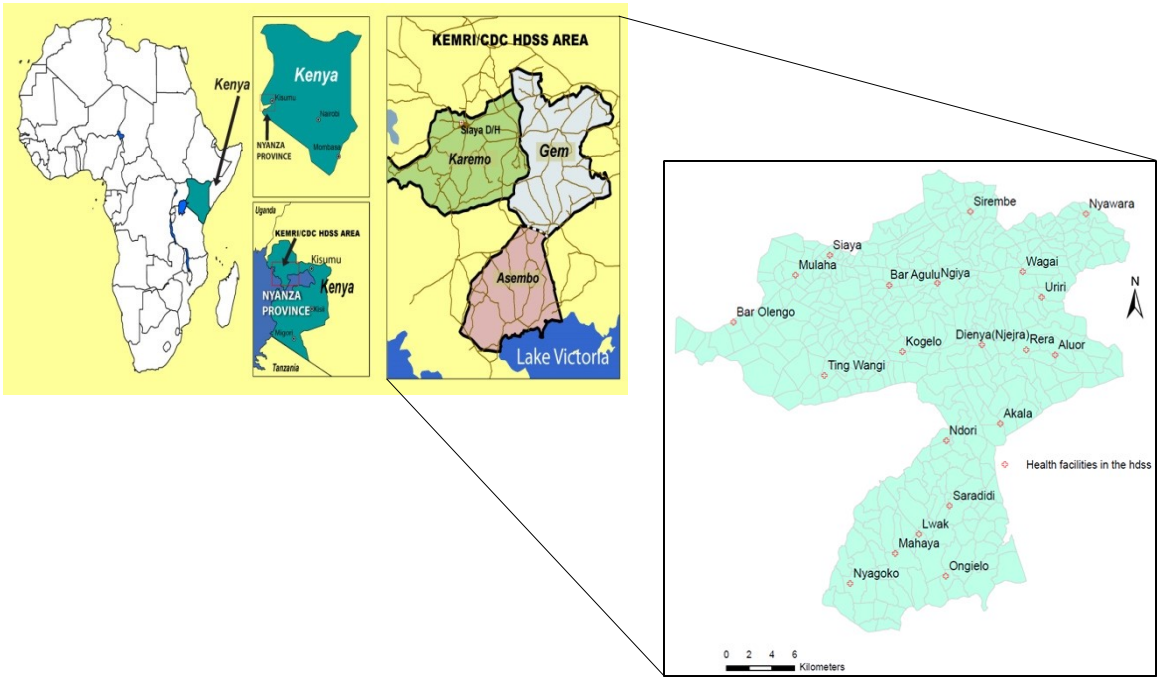


Figure 2.1: KHDSS study area in western Kenya showing villages and selected health facilities

2.2.2. Malaria incidence

From 2007 to 2012, in the KHDSS region encompassing 385 villages, children aged 5 and under who were admitted to medical facilities with confirmed or suspected fevers underwent malaria testing. The test involved taking a blood drop, placing it on a slide, staining it with

Giemsa solution, and examining it under a microscope. A confirmed case of malaria was identified if parasites were present along with a reported or documented fever.

To ensure data quality and precision, microscopists specialized in blood slide reading were assessed for their proficiency every three months. Additionally, 10% of the slides from each testing site were re-examined as a quality check. Patients from the KHDSS area were pinpointed using unique personal identification numbers, and their data linked to their household locations. The analysis included only those children registered as residents of the KHDSS region.

To measure the malaria incidence rate, we divided the number of confirmed cases by the total amount of time contributed by children under 5, both aggregated by village and month. The resulting data, monthly count of malaria-positive cases per village, served as our outcome variable. We also identified seasonal variations in malaria transmission, classifying "low peaks" in February and August and "high peaks" from January, March to July, and September to December, when transmission rates were significantly higher.

2.2.3. Climatic and environmental factors

We extracted remotely sensed climatic and environmental data collected by satellites and stored as rasters based on geocoded locations of the compounds from which participants admitted at local hospitals come from. Day land surface temperature (LST) at daily temporal and 1000m spatial resolution; Enhanced vegetation index (EVI) an advanced vegetation index created with higher sensitivity to biomass, atmospheric background, and soil condition at 16 days temporal and 500m spatial resolution; Land cover (LC) the observed physical cover and natural use of the land including the vegetation and human construction over the earth's surface at yearly temporal and 500m spatial resolution; were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) website (16). Decadal rainfall data were downloaded via the United States Geological Surveys' Famine Early Warning Systems Network (USGS FEWS)

data portal (https://earlywarning.usgs.gov/fews); while altitude data was extracted from the Shuttle Radar Topography Mission (SRTM) by the U.S. Geological Survey - Earth Resources Observation and Science (USGS EROS) Data Center (https://eros.usgs.gov/elevation-products).

To account for the climatic/environmental lag effects on malaria transmission, we extracted the variables up to three months prior to the month of illness for each household from which a case came from. We categorized temperature data into three levels as $>25^0$, 25^0 -35^0, $>35^0$, based on the nonlinear effect of temperature on mosquito and parasite survival and hence malaria transmission (Bayoh & Lindsay, 2003; Snow & Omumbo, 2006). Land cover was included as a continuous variable based on the year of data collection while EVI was also included as a continuous variable by month. Rainfall was classified into 3 categories as; greater than 28mm, between 28mm and 50mm, and more than 50mm; based on the biological plausibility of mosquito/parasite life cycle and malaria infection in humans (Rotejanaprasert et al., 2021). The effects were measured by looking at different lags or time periods. The first month's effect was referred to as "lag0", the effect from the previous month as "lag_1", the effect from two months prior as "lag_2", and the effect from three months prior as "lag_3". These effects were considered for each of the other environmental and climatic factors, including Land Surface Temperature (LST), Enhanced Vegetation Index (EVI), and Rainfall.

We generated average values over different periods. The two-month average included the current month and one month prior to collection; the three-month average consisted of the current month and the two preceding months; and the four-month average incorporated the current month along with the three previous months prior to collection.

2.2.4. Household and socioeconomic data

From the KHDSS household surveys, data on changes in migration status, deaths and births were used to compute person-time in years under observation (pyo) for all children aged 0-59 months who met the KHDSS residency requirements. The pyo was summed from 1st January 2007 or the date of enrolment, until the child exited or lost residency status due to death, out-migration, loss to follow-up, or reaching the observation period's end on December 31, 2012. Person-time was adjusted by a factor of 0.5 to account for health-seeking behaviour in case of fever reported at home among children in this study area (Bigogo et al., 2010).

In addition, household level data on socio-economic indicators, house types, and bed net ownership were obtained. Using these data, we created a composite socio-economic status index through multiple correspondence analysis (MCA) as previously described (Amek et al., 2015). Household' scores were aggregated at village level and ranked into five quintiles, classified as "Richest" for the well off, "Richer", "Middle", "Poorer" and "Poorest" for the lowest rank.

Bed net ownership was calculated as the percentage of households per village owning at least one net for every two people (Roll Back Malaria Monitoring & Evaluation Reference Group, 2018). Lastly, the distance to health facilities was determined by measuring the Euclidean difference between the household and the nearest health facility in kilometres. These distances were aggregated at the village level then classified into 3 categories as; less than 1km, 1 to 2km, and greater than 2km.

2.2.5. Bayesian modelling

Using unique Personal Identification Numbers (PIDs), data from the health facilities were linked to households, and Remote Sensing (RS) data. This was the aggregated at monthly and village level. We calculated malaria incidence rates by dividing the number of microscopically

confirmed positives slide over the under-5 village population in that month. The incidence rate was further adjusted by multiplying the person time of years (pyo) contributed by children under five years by a factor of 0.5 to account for the health-seeking behaviour in this age group as shown in previous studies (Bigogo et al., 2010).

To explore the factors influencing malaria incidence, Bayesian spatiotemporal negative binomial models were developed, adjusting for; bed net use, seasonality, peak months of transmission, climatic/environmental factors, distance to health facilities and water bodies, socioeconomic status, as well as spatial and temporal random effects. Likewise, the pyo was treated as an offset term in the model. An initial investigation of the association between the covariates and incidence rate was carried out in Stata 14 for variable selection. Only those variables that were statistically significant at a 95% confidence interval were included in the Bayesian models. Two models were constructed; a non-spatial model including only the aforementioned covariates; and a second model that took into consideration the spatial and temporal effects. The spatial variation was treated as village-specific random effects, with latent observations of a spatial Gaussian process with a mean of zero and a covariance assuming an exponential variation function of distance between two villages (Diggle et al., 1998). Temporal variation was modelled using a first-order autoregressive process with monthly random effects. Suitability of the models was checked using the Deviance Information Criteria (DIC), where the best model was determined by the smallest DIC value (Spiegelhalter et al., 2002). The Bayesian models were fitted in OpenBugs version 3.1.2 (Imperial College and Medical Research Council London, UK) using Markov Chain Monte Carlo (MCMC) simulation.

Ethical statement

Written informed consent was obtained from the heads of all compounds within the KHDSS, who agreed to participate in the studies. They provided consent on behalf of all members, including children, for the collection of household data. In instances where individuals from the

KHDSS, or caregivers in the case of minors, visited health facilities for medical care, additional consent was obtained for the collection and use of data gathered during the visit. The protocol for the KHDSS (within which this study is incorporated) and the consent procedures undergo annual review and approval by the ethical review boards of both KEMRI (#1801, Nairobi, Kenya) and CDC (#3308, Atlanta, GA).

2.3. **Results**

Between January 2007 and December 2018, 15,095 children under the age of 5 years were admitted to select sentinel health facilities in the KHDSS. During this period, children under 5 years contributed 68,240 person-years of observation (pyo) for a malaria incidence rate of 628 cases per 1,000 pyo. Annually, malaria incidence ranged from 448 cases per 1,000 pyo in 2007, peaked at 775 cases per 1,000 pyo in 2009, and gradually declined to 540 cases per 1,000 pyo by 2012 (Table 2.1).

Table 2.1: Distribution of average malaria incidence rates by year of study

Year	Person years of observation (pyo)	Malaria cases	Adj. Malaria incidence rate (cases per 1000 pyo)
2007	9571	4290	448 (435-462)
2008	14963	10160	679 (666-692)
2009	11736	9099	775 (759-791)
2010	11975	8343	697 (682-712)
2011	10523	5826	554 (540-568)
2012	9472	5115	540 (525-555)
Total	**68240**	**42833**	**628 (622-634)**

From the overall surveys, the spatial distribution of the cases in all the 370 villages is shown below (Figure 2.2).

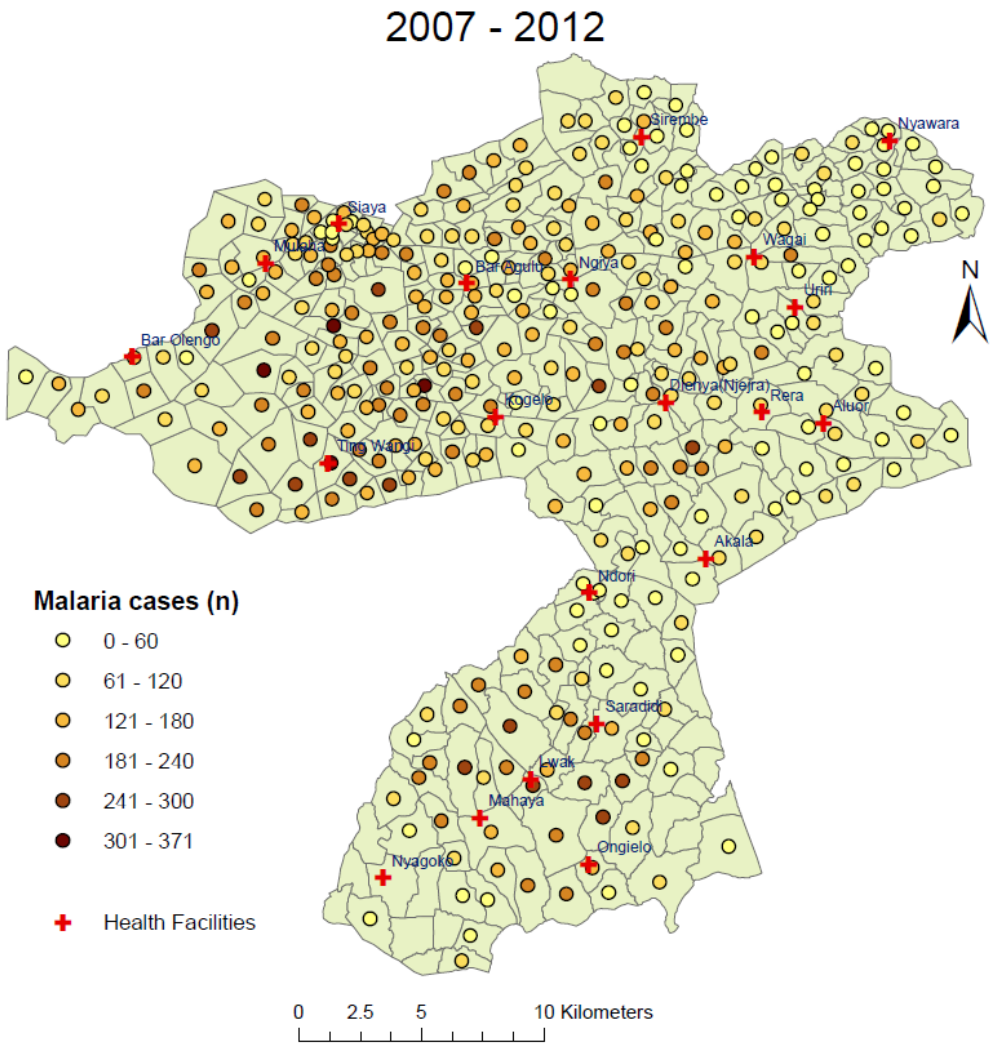


Figure 2.2: Observed malaria cases at health facilities for the period 2007-2012

The distribution of cases by month ranged between 344 cases per 1000 pyo in February 2007 to a maximum of 838 cases per 1000 pyo in December 2009. There were observable fluctuations in incidence rates over the course of the year, exhibiting distinct peaks and troughs. Specifically, these peaks occurred in two periods: Between May to July and November to January of each year (Figure 2.3). The temporal association between malaria incidence and selected environmental or climatic variables is also shown below. Notably, the peaks in monthly Enhanced Vegetation Index (EVI) closely matched with the peaks of malaria incidence rates, while increases in temperature and rainfall showed a lagged effect in relation to incidence. Overall, the temporal trend in incidence increased between 2007 and 2009, followed by a gradual decline in the subsequent months.

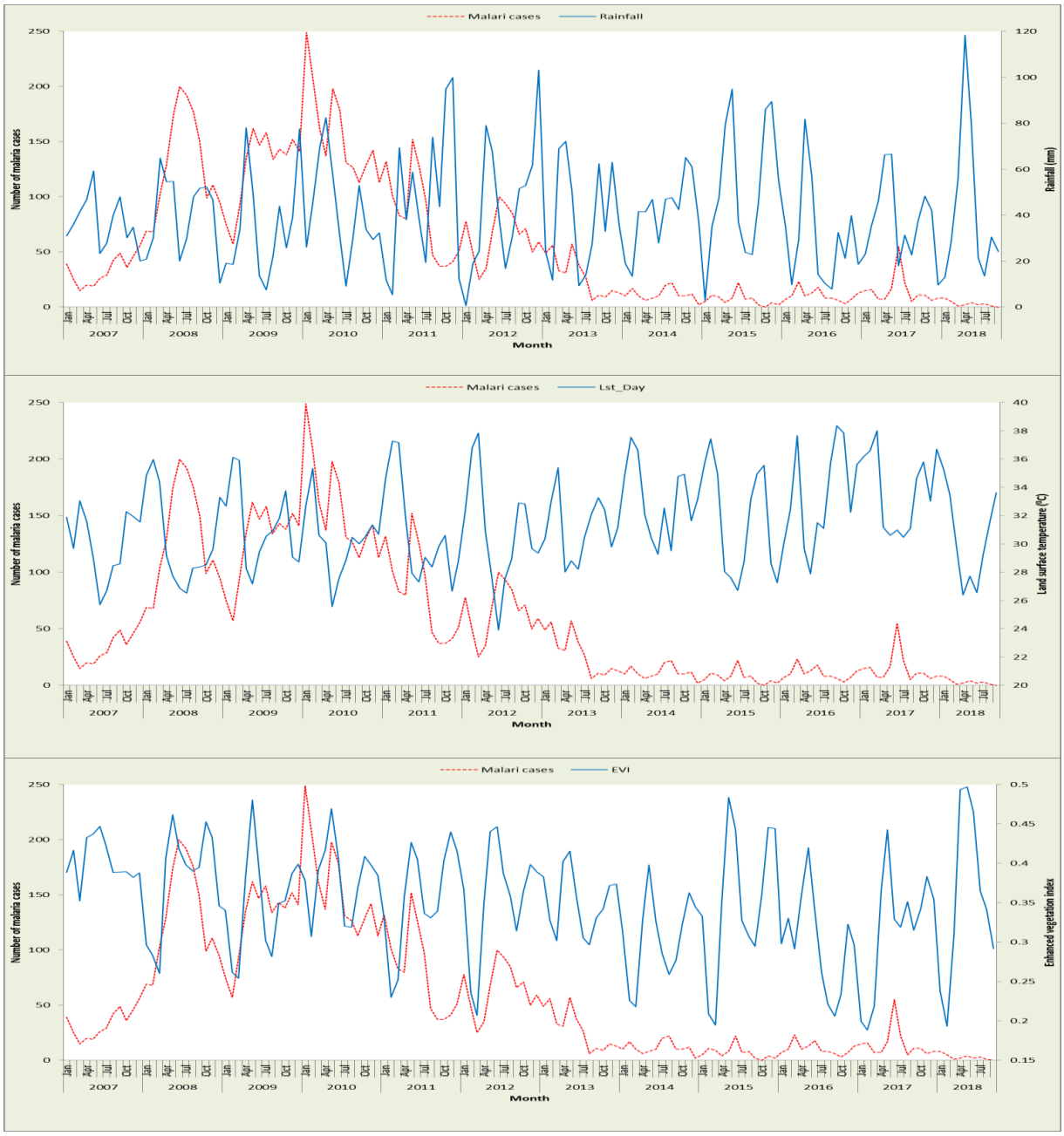


Figure 2.3: Temporal variation of climatic/environmental and factors malaria incidence in children under 5 years of age

Model based results

We fitted two distinct Bayesian models to assess the effects on malaria incidence: model 1, excluding spatial and temporal factors, and model 2, incorporating both spatial and temporal effects. Table 2.2 illustrates the adjusted parameter estimates along with their corresponding 95% Bayesian credible intervals (BCI) for malaria incidence.

In the non-spatial model (model 1), several factors such as proximity to health facilities, socio-economic status (ses), the year of the study, study area, transmission peaks, altitude, land cover, and EVI were identified as statistically important in influencing malaria incidence among children under 5. Conversely, in the spatio-temporal model (model 2), only socio-economic status, study area, transmission peaks, and altitude emerged as important factors.

Both models pointed to the study area, altitude, and peak months of transmission as the most potent risk factors. Specifically, the Karemo region exhibited a three-fold increase (IRR=3.19; 95% CI: 2.05-5.68) in malaria incidence relative to the Asembo region. Although transmission persists throughout the year, the risk was 1.3 times higher during the peak months of March to July and September to January (Figure 2.3), compared to February and August (IRR=1.32; 95% CI: 1.20-1.47). Additionally, higher altitude areas were associated with a reduced risk of infection as opposed to lower-altitude regions.

The superior model was found to be the spatio-temporal Bayesian (model 2), as it demonstrated a lower DIC value. Furthermore, the estimated spatial variation was strongly associated to malaria incidence, overshadowing temporal correlations. The minimum distance at which a 5% significant spatial correlation was detected stood at 10.4 km (95% CI: 8.08-32.83).

Table 2.2: Spatio-temporal Posterior estimates (median) of predictors for malaria incidence among children aged >5 years

Covariates		Non spatio-temporal model IRR (95% BCI)	Spatio-temporal model IRR (95% BCI)
Distance to health facility	0 – 1 km	1	1
	1 – 2 km	1.15 (1.10-1.21)	1.05 (0.75-1.35)
	>2 km	1.43 (1.35-1.51)	1.22 (0.80-1.55)
Ses	Poorest	1	1
	Poorer	0.97 (0.91-1.03)	0.92 (0.86-0.97)
	Middle	0.97 (0.91-1.03)	1.01 (0.96-1.07)
	Richer	0.93 (0.88-0.99)	0.99 (0.93-1.04)
	Richest	0.87 (0.82-0.93)	1.02 (0.96-1.08)
Year	2007	1	1
	2008	1.61 (1.50-1.73)	1.47 (0.99-2.55)
	2009	1.76 (1.62-1.89)	1.40 (0.79-2.49)
	2010	1.67 (1.56-1.79)	1.26 (0.62-2.77)
	2011	1.30 (1.20-1.40)	0.86 (0.44-2.17)
	2012	1.27 (1.17-1.37)	0.80 (0.39-1.96)
Area	Asembo	1	1
	Gem	1.56 (1.45-1.69)	1.19 (0.60-3.11)
	Karemo	1.76 (1.65-1.88)	3.19 (2.05-5.68)
Transmission season	Low	1	1
	High	1.33 (1.26-1.41)	1.32 (1.20-1.47)
Altitude	1147 – 1243	1	1
	1244 – 1293	0.89 (0.84-0.94)	0.62 (0.41-1.01)
	1294 – 1327	0.68 (0.64-0.72)	0.45 (0.25-0.76)
	1328 – 1365	0.82 (0.76-0.87)	0.74 (0.30-1.15)
	>1365	0.58 (0.54-0.63)	0.69 (0.25-1.24)
Land cover	Grassland/Savannah	1	1
	Permanent Wetlands	1.08 (0.78-1.42)	0.88 (0.64-1.19)
	Cropland	1.84 (1.35-2.35)	1.29 (0.97-1.73)
	Built up/Urban	1.63 (1.20-2.08)	1.24 (0.93-1.66)
	Natural Vegetation	1.52 (1.13-1.93)	1.20 (0.90-1.62)
EVI (Lag0)		1.34 (1.04-1.71)	1.00 (0.70-1.67)
Rainfall (Lag_3)	<28mm	1	1
	>28mm – 50mm	1.00 (0.96-1.05)	0.99 (0.93-1.05)
	>50mm	0.99 (0.94-1.04)	0.96 (0.88-1.04)
LST (Day)	$<25^0$c	1	1
	25^0c – 35^0c	1.01 (0.94-1.07)	1.00 (0.93-1.08)
	$>35^0$c	0.95 (0.84-1.07)	0.88 (0.77-1.01)
Spatial Variation		-	381.4 (267.9-1204.0)
Temporal Variation		-	0.04 (0.03-0.06)
Range**		-	10.4 (8.08-32.83)
DIC		77570.0	63860.0

BCI=Bayesian credible interval
DIC=deviance information criterion
** Minimum distance in kilometres at which spatial correlation is significant at 5%

2.4. Discussion

In this study, we provide robust estimates of malaria incidence at population level in children under five years in a malaria endemic region of western Kenya between 2007 and 2012. The incidence was relatively high compared to other regions of the country and remained stable during the whole study period. Comprising 15,095 young children admitted to designated health centres, this study builds a robust foundation for examining trends, contributing factors, and spatiotemporal links related to malaria occurrences in this at-risk group. The outcomes of this study elucidate a complex landscape of malaria prevalence, highlighting key findings and suggesting possible areas of improvement in malaria incidence estimation. These have been revised and explored further in guiding follow-up analyses.

Findings this study also show that incidence among children aged five years and below was extremely high compared to other regional and countrywide estimates. The estimated incidence rates of between 450 to 775 cases per 1,000 pyo were way above the regional average of less than 100 cases per 1000 pyo given by the WHO/World Bank (Roser & Ritchie, 2019) and national indicator surveys during the same period (Division of Malaria Control [Ministry of Public Health and Sanitation] et al., 2011). This reinforces the need for continued and intense localized control measures to reduce the burden of malaria in the lake endemic region, different from other places in the country (Division of National Malaria Programme (DNMP) & ICF, 2021).

While the rate of transmission remained elevated year-round, our data point to notable seasonality. Transmission peaks were consistently experienced between March to July and September to January every year. This confirms that malaria is an enduring public health challenge that requires sustained monitoring and control measures. Our recent subsequent analysis confirmed this finding (Nyawanda et al., 2023), and had been observed in other studies previously (Otambo et al., 2022; Sewe et al., 2015). This warrants having control and preventative interventions to be enhanced during and following the peaks in the rainy seasons.

Interestingly, climatic elements did not serve as decisive drivers of malaria prevalence in this analysis. Although mean temperature and rainfall exhibited some association with the incidence rates, the relationship was not statistically important. Lagged peaks in mean monthly temperature and rainfall rose just before the peaks in monthly malaria incidence rates. A previous study in the same region on a less fine scale using aggregated data at the district level (Sewe et al., 2015) and subsequent analysis of a smaller subset in this same population that was followed-up more actively (Nyawanda et al., 2023) showed lagged effects of both rain and day temperatures. In that sub-analysis, increase in mean temperature was associated with a decline in malaria incidence while rainfall increase was associated with lower incidence. The lack of association could have been due to the surveillance in the larger population being more passive compared to the smaller intense region with an active surveillance. While the aggregated data might have masked the spatial differences in our study. Alternatively, this could have been because the study period was bit too short for this kind of analysis since subsequent sub-analysis for double the period of time showed otherwise (Nyawanda et al., 2023).

Enhanced vegetation index (EVI), a short term vegetation indicator of a combination of suitable climatic conditions of optimal rainfall and temperature (Wang et al., 2022) in the previous few weeks was associated with increased incidence. The correlation between EVI and incidence rates reveals a close relationship between vegetation and malaria transmission. This could indicate that places with denser vegetation provide optimal breeding grounds for mosquitoes that carry malaria parasites (Hinne et al., 2021; Obsomer et al., 2007). This association confirms that a combination of suitable climatic conditions are important drivers of malaria transmission, rather than just one distinct variable (Nyawanda et al., 2023).

Socio-economic status, study area, and altitude emerged as significant risk factors. These determinants align with other research, suggesting that malaria risk in the Karemo region was higher than in the Asembo and Gem regions (Khagayi et al., 2017) which could be due to other health systems indicators not measure but vary in the three sub-regions. The risk diminished in

areas of higher altitude and been observed in other studies too (Afrane et al., 2014; Oduma et al., 2023). Lower temperatures at higher altitudes, which directly affect development and survival of mosquitos and Plasmodium parasites (Githeko et al., 2006; Villena et al., 2022), might have caused the observed pattern. These variables hint at an intricate interplay among socio-economic factors, geography, and healthcare-seeking behaviours, which might be more significant due to the limited geographical scope of the study.

Some limitations of the study could have been due to incorrectly adjusting for health seeking behaviour by village and month since the adjustment used was a one-time survey in the KHDSS and assumed to be uniform throughout the study period and by month. This adjustment was also based on children seeking hospital care for reported fever at home and not confirmed malaria cases. Our subsequent in-depth analysis of a smaller, actively monitored subgroup from the same population, however, accounted for these shortcomings.

2.5. **Conclusion**

Robust surveillance remains a cornerstone in the battle against malaria, particularly given the much higher incidence rates observed in this study compared to regional estimates. The evident spatial and seasonal variations, regional and socioeconomic influences stress the necessity for localized, timely preventive and control strategies.

Declarations

Ethical consideration

The study protocol for the KHDSS the household surveys was approved by the Kenya Medical Research Institute (#1801) and Centers for Disease Control and Prevention (#3308) institutional review boards. Informed written consent was obtained from the compounds heads, while additional consent was sought from individuals or the caretakers of those under 18 years admitted to the different health facilities.

Data availability

Data for this study was obtained from the KHDSS with the approval of its steering committee. For data requests or inquiries, please contact the Steering Committee via Dr. Stephen Munga. (Smunga@kemri.org).

Competing interests

The authors declare that there are no competing interests.

Funding

This work was supported by the European Research Council (ERC) IMCCA grant number 323180 and the Swiss National Foundation (SNF) program for Research on Global Issues for Development (R4D) project number IZ01Z0-147286.

Authors' contributions

SK designed the initial study, carried out the surveillance, conceived the study, analysed the data and drafted the initial manuscript; PV conceived the study and supported the data analysis; BN and AB supported data analysis and interpretation; all authors gave intellectual inputs in analysis, revised the drafts, read and approved the final manuscript.

Acknowledgement

We extend our heartfelt thanks to the HDSS community in Siaya County for their enduring support and active participation in our surveillance efforts over the years. We are also grateful to the past and current staff of the KHDSS and the Malaria Branch at KEMRI-CGHR for their invaluable assistance in managing the platforms and handling data collection and management. This article has been published with the permission of the Director of the Kenya Medical Research Institute.

2.6. Appendix

Appendix 2.6.1: Bayesian models and methods

In our study, we employ Bayesian models and methods to analyse malaria cases among children under 5 years of age. Let $\mathbf{Y}_{jt}$ represent the average number of malaria cases in village *j* at time interval *t*. We assume $\mathbf{Y}_{jt}$ follows a negative binomial distribution:

$$\mathbf{Y}_{jt} \sim \text{NegBin}\ (\mathbf{P}_{jt}, \boldsymbol{r})$$

Where, P_{jt}, is the proportion of deaths in village *j* at time interval t and *r* is the dispersion parameter with,

$$\mu_j = r\frac{1-p}{p} \text{ and } \sigma_i^2 = r(1-p)p^{-2}.$$

We model the relationship between various covariates *X* and mortality status in each village using a logit model:

$$logit(\mu_{jt}) = logit(N_{jt}) + \beta_0 + \sum_1^k \beta_u X_u + \phi_j + \varepsilon_t \qquad , u=1,2,\ldots\ldots k$$

Here, μ_{jt} represents the number of deaths in village *j* at time *t*, while N_{jt} denotes the total person-time in discrete months. β_i are the regression coefficients, ϕ_j signifies village specific spatial effects and ε_t represents temporal (monthly) random effects.

We assumed ϕ_j to be a parameter from a latent spatial process modeled by a Gaussian distribution with an exponential correlation function based on distance between villages irrespective of the direction using an exponential correlation function.

$$\underline{\phi} \sim \text{MVN}(0, \Sigma)\,,\ \Sigma_{kl} = \sigma_1^2 \exp(-\rho d_{kl})$$

Where σ_1^2 is the spatial variation, d_{kl} is the distance between villages k and l, and ρ is the rate of correlation decay with increasing distance. The minimum distance at which the spatial variation is less than 5% (range) and was obtained from the value $3/\rho$. The temporal effect (ε_t) is modelled by an autoregressive process of order 2.

We employed non-informative normal prior distributions for the regression coefficients with mean zero and large variance for the β_i i=1,……; inverse gamma priors for $r \sim IG(1.01, 0.001)$, σ_e^2 and σ^2; and a gamma prior for ρ, i,e. $\sigma_e^2, \sigma^2 \sim IG(2.01, 1.01)$ and $\rho \sim G(0.01, 0.01)$.

The model was implemented using Markov Chain Monte Carlo (MCMC) simulations in OpenBUGS version 3.1.2. (Imperial College and Medical Council, London, UK) to estimate model parameters (Gelfand & Smith, 1990). We ran a single chain sampler and discarded the initial 5,000 iterations. Convergence was assessed after 100,000 iterations using the Gelman-Rubin diagnostics. (Gelman & Rubin, 1992).

Appendix 2.6.2: Posterior inclusion probabilities for climatic covariates

Indicator	Probability of inclusion (%)
Climatic factors	
Rainfall	
Rain_01	0.0
Rain_01*	0.0
Rain_012	0.0
Rain_012*	0.0
Rain_0123	**100**
Rain_0123*	0.0
EVI	
EVI_01	**100**
EVI_01*	0.0
EVI_012	0.0
EVI_012*	0.0
EVI_0123	0.0
EVI_0123*	0.0
LSTD	
LSTD_01	19.0
LSTD_01*	0.0
LSTD_012	**58.0**
LSTD_012*	0.0
LSTD_0123	21.0
LSTD_0123*	0.0
LSTN	
LSTN_01	0.0
LSTN_01*	0.0
LSTN_012	9.8
LSTN_012*	0.0
LSTN_0123	**79.4**
LSTN_0123*	0.0
Altitude	
Altitude	0.0
Altitude*	**100.0**

**Categorical*

In bold: variables with highest inclusion probability that were fitted in the spatio-temporal model

3. Chapter 3: Bayesian spatio-temporal modelling of mortality in relation to malaria incidence in western Kenya

Sammy Khagayi[1,2,3], Niagara Amek[1], Godfrey Bigogo[1], Frank Odhiambo[1] and Penelope Vounatsou[2,3§].

1. Kenya Medical Research Institute-Center for Global Health Research, Kisumu, Kenya
2. Swiss Tropical and Public Health Institute, Basel, Switzerland
3. University of Basel, Switzerland

[§]Corresponding author

This paper has been published in *PLOS One* July 2017

Abstract

Introduction:
The effect of malaria exposure on mortality using health facility incidence data as a measure of transmission has not been well investigated. Health and demographic surveillance systems (HDSS) routinely capture data on mortality, interventions and other household related indicators, offering a unique platform for estimating and monitoring the incidence-mortality relationship in space and time.

Methods:
Mortality data from the HDSS located in Western Kenya collected from 2007 to 2012 and linked to health facility incidence data were analysed using Bayesian spatio-temporal survival models to investigate the relation between mortality (all-cause/malaria-specific) and malaria incidence across all age groups. The analysis adjusted for insecticide-treated net (ITN) ownership, socio-economic status (SES), distance to health facilities and altitude. The estimates obtained were used to quantify excess mortality due to malaria exposure.

Results:
Our models identified a strong positive relationship between slide positivity rate (SPR) and all-cause mortality in young children 1-4 years (HR=4.29; 95% CI: 2.78-13.29) and all ages combined (HR=1.55; 1.04-2.80). SPR had a strong positive association with malaria-specific mortality in young children (HR=9.48; 5.11-37.94), however, in older children (5-14 years), it was associated with a reduction in malaria specific mortality (HR=0.02; 0.003-0.33).

Conclusion:
SPR as a measure of transmission captures well the association between malaria transmission intensity and all-cause/malaria mortality. This offers a quick and efficient way to monitor malaria burden. Excess mortality estimates indicate that small changes in malaria incidence substantially reduce overall and malaria specific mortality.

Key Words: Malaria mortality, Incidence, Bayesian spatio-temporal, Health and demographic surveillance system, Western Kenya

3.1. Introduction

Morbidity and mortality estimates over the last decade across age groups in sub Saharan Africa (SSA) remain high compared to other regions despite an overall global reduction. The biggest burden is due to infectious diseases that largely affect children below 5 years of age with one of the main drivers of these consistently high estimates being malaria (Murray et al., 2014; World Health Organization, 2015b). Recent studies and estimates show that malaria in SSA has reduced considerably, with a drop of over 37 % for cases and 60 % of deaths between the years 2000 and 2015 (Bhatt et al., 2015). In western Kenya, the Kenya Medical Research Institute and Centers for Disease Control and Prevention's run a Health and Demographic Surveillance System (KHDSS) which has shown that between 2003 and 2010 there was a 67 % reduction in malaria mortality in all ages and 70 % in children below the age of 5 years even though it remains a leading cause of death (Desai et al., 2014).

Malaria infection is driven by different factors and measuring its burden has largely been problematic, especially in areas where the disease's health impact is huge. Malaria transmission intensity is an important measure of this burden and can largely be classified into infection in humans (parasite prevalence and incidence rates), interaction between mosquitos and humans (entomological inoculation rate (EIR)) and vector measures like (mosquito density, vectoral capacity and sporozoite rate). The links between the above drivers of transmission and mortality still need further investigation so as to understand and quantify them well. Recent studies have concentrated on measures of vector density and interaction between mosquitos and humans to measure transmission intensity and link it to both all-cause and malaria specific mortality (Amek, 2013; Carneiro et al., 2010; Gething et al., 2016; Rumisha et al., 2014). Meanwhile measures of infection in humans have largely been carried out either through parasite prevalence or disease rates as these can be done as one-off surveys with ease especially during peak transmission times. Incidence as a measure of transmission has largely been unused since it requires more investment of time and resources. Ensuring the correct denominator is used

while taking into account health seeking behaviour, poor health systems and diagnostic challenges limit its usability and makes quality data unavailable.

Slide positivity rate (SPR), has been used as a surrogate for malaria incidence and a key indicator of temporal trends in malaria disease burden. It measures the proportion of slide confirmed malaria positive cases out of cases examined, and has been shown to be a good predictor of incidence (Bi et al., 2012; Boyce et al., 2016; Jensen et al., 2009). It is relatively inexpensive compared to other measures of transmission and easy to monitor at sentinel health facilities hence a useful measure of not only trends but the overall burden of malaria disease.

Despite global improvements in data collection and use, there is a large gap of adequate and quality data on malaria cases and deaths in low and middle income countries. Estimates provided from these data are fraught with inherent differences in data collection methodology, analysis and interpretations. This could be attributed to poor health surveillance systems that are incapable of collecting quality data to be used in informing policy (Mikkelsen et al., 2015; Setel et al., 2007). It is important therefore to explore and determine the best modes of appropriately measuring malaria burden so as to allow for accurate determination of progress and assess the contribution of different interventions especially at local levels. With this in mind, health and demographic surveillance system (HDSS) sites were established in different regions of SSA, Asia and Oceania to supplement efforts in providing accurate data on demography and public health (Adazu et al., 2005; Ye et al., 2012). These HDSS sites have accumulated a wealth of information that is well aligned in both space and time, offering a unique platform to monitor and provide accurate estimates of disease at a local level. The other benefit that HDSS sites confers is the intensity and longevity of its operations which provide consistently regular longitudinal data that can aid policy making in resource strained settings (Bos, 2004). In addition to household survey data, HDSS sites, using verbal autopsy (VA) collect data on morality that is used to determine cause of death, which is an important tool,

especially where conventional autopsy data is not available (Leitao et al., 2013). HDSS sites are therefore an important platform for malaria burden estimation.

The Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) project sought to analyse data on malaria transmission intensity and burden in Africa from 2002 to 2004. Project results indicated that malaria transmission as measured by EIR drives mortality especially in younger children and showed the importance of small scale heterogeneity in estimating malaria related deaths. However their analysis did not include data on malaria morbidity or interventions. Furthermore, the two year duration of the project could not allow adequate assessment of trends and changes in malaria mortality in relation to interventions.

In this study, malaria indicator incidence data collected from selected health facilities was linked to household data so as to improve estimates of malaria related burden by including more malaria transmission indicators over a period of 6 years in the Kenya Medical Research Institute's Health and Demographic Surveillance System (KHDSS) of western Kenya. We also sought to explore the relationship between slide positivity rate from these health facilities as a measure of incidence and all-cause/malaria-specific mortality derived from verbal autopsy (VA). Bayesian geostatistical spatio-temporal models were used to estimate the contribution of SPR to malaria mortality in an area of almost year round transmission.

3.2. Materials and methods

3.2.1. Study area and population

The KHDSS has been described in details elsewhere (Adazu et al., 2005; Odhiambo et al., 2012). In brief, the KHDSS is located in three regions (Asembo, Gem and Karemo) of rural western Kenya, Siaya County (Figure 3.1). The study area borders the northern shores of Lake Victoria; is spread over 700km^2 divided into 385 villages with a mid-year population of over 240,600 people in 58,700 households. The population is predominantly one ethnic group who

derive a big part of their livelihood from subsistence farming and fishing. It is located in a malaria endemic zone having round the year transmission with peaks in May-June and November-December (Amek, 2013). Infectious diseases and HIV/AIDS are the other important causes of morbidity and mortality (Adazu et al., 2005; Division of Malaria Control [Ministry of Public Health and Sanitation] et al., 2011; Odhiambo et al., 2012).

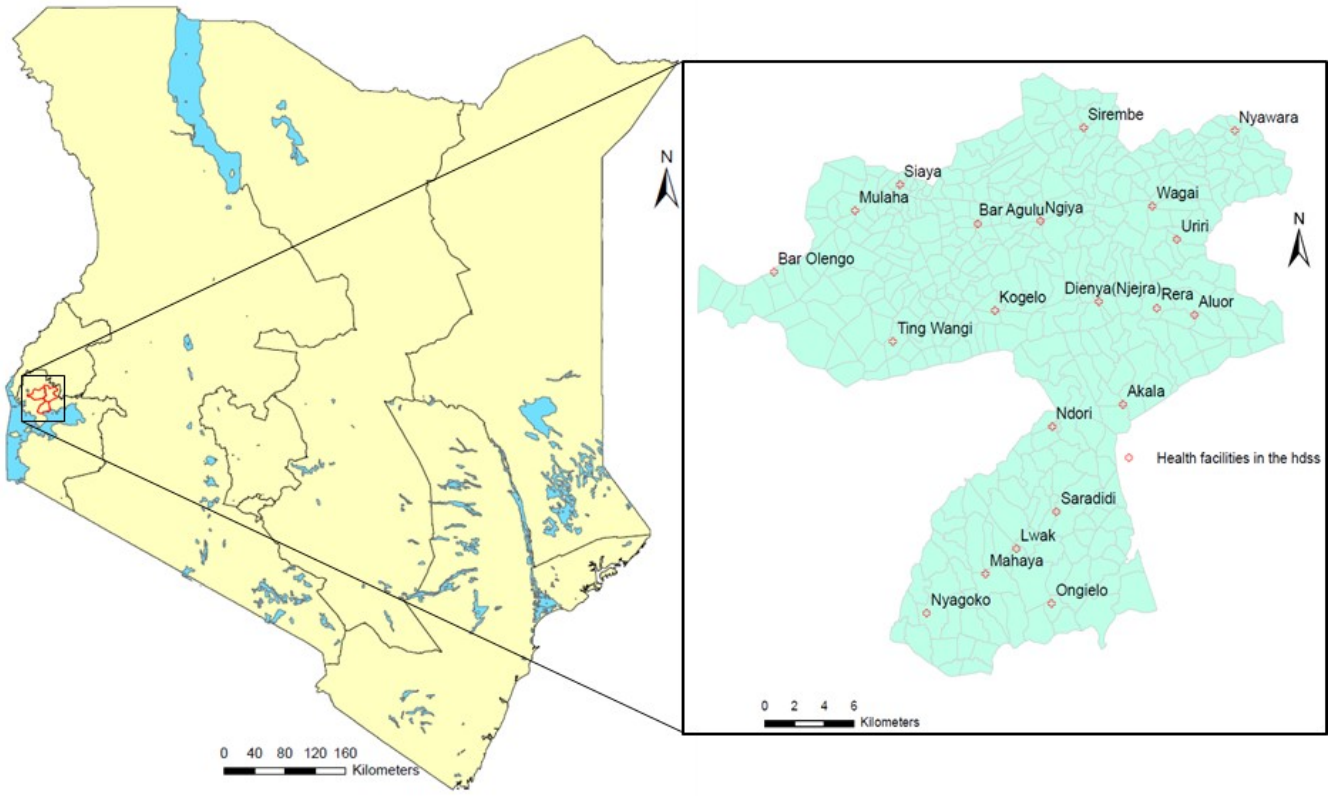


Figure 3.1: KHDSS study area in western Kenya showing villages and sentinel health facilities

3.2.2. Slide positivity rates

Children aged below 14 years admitted at selected sentinel health facilities within the KHDSS (Figure 3.1) from January 2007 to December 2012 with history of or documented fever were screened for malaria infection by collecting a blood slide for microscopy examination regardless of their residency status. Blood smears were collected and screened microscopically after staining with giemsa. Presence of parasites and reported or documented fever was classified as malaria infection. All the microscopists at these health facilities are evaluated after every 3 months for competency in reading malaria microscopy slides and 10 % of the slides are

rechecked for quality control. However for this analysis, we restricted our numbers to patients registered and residing in the study area to calculate the slide positivity rate (SPR). The data were aggregated at monthly intervals and the slide positivity rate (SPR) calculated as the proportion of malaria positive cases as determined by microscopy out of the total cases tested.

3.2.3. KHDSS surveillance data and verbal autopsy

The KHDSS population has been monitored from an initial census in 2001at household level to establish the number of people in a geographically defined region followed by tri-annual visits after every 4 months to establish any changes due to migration, deaths or births. In addition, other data was collected at household level for socio-economic indicators, house types and insecticide-treated net (ITN) ownership.

The verbal autopsy (VA) methodology and data used in this study has been described in details elsewhere (Amek et al., 2014). After death of a registered KHDSS resident, a notification is filled as soon as possible by a reporter based in the same village. At least 3 weeks are given to allow for mourning before an interview is conducted. A detailed questionnaire is printed to collect data on the deceased last disease, signs, symptoms and medical history. A trained interviewer looks for the most appropriate interviewee who was closest to the deceased and knew about the illness, disease or condition that led to death to administer the questionnaire.

The data is then captured to an electronic database, validated and processed using the InterVA program which is a computer based expert opinion algorithm that is based on the Bayes theorem. It uses probabilistic methods to interpret verbal autopsy data using *a priori* approximations of probabilities related to diseases and last symptoms exhibited by the diseased so as to determine the most probable cause of death based on these data (Byass et al., 2012). Output from the InterVA model was used to determine the most probable causes of death including malaria.

Data on household assets ownership and house characteristics was used to create a composite socio-economic status index using the multiple correspondence analysis (MCA) technique as described in previous studies (Amek et al., 2015). The household scores were then aggregated at village level and ranked into 3 categories, i.e. least poor for the well off, poor for the average and poorest for the lowest rank. Bed net ownership was used as a measure of existing interventions that have implemented in the area. It was calculated as a percentage of households owning at least one net for every two people per village (MEASURE Evaluation et al., 2013). The data used are described in details in appendix 1.

3.2.4. Statistical analysis

All registered residents in the KHDSS villages during the period 2007 to 2012 were included in the analysis. To qualify as a resident, a person has to stay continuously in the study area for at least 3 calendar months or is born to a resident mother while she is a resident. The participants were grouped into 6 age groups as follows: 0-28 days (neonates), 1-11 months (infants), 1-4 years (child), 5-14 years (older child), 15-59 (adults) and 60+ (elderly).

Person time at risk in months was calculated as the total time spent in the study area from date of enrolment until they exited through outmigration or death; alternatively they stopped being observed due to loss to follow-up or reached the end of the observation period set at 31st December 2012.

Crude and age specific all-cause mortality and malaria rates were calculated by dividing the deaths in each group with the total person years observed (pyo). For each age group, Cox proportional hazards models were approximated using negative binomial regression (Manda and Meyer 2005) with time to death in person months as discrete contribution of each resident summed up at village level. Initial exploratory analysis was carried out in Stata 14 (Stata Corporation USA) to assess bivariate models of the relationship between the different factors and all-cause or malaria specific mortality. We included all the variables that were at 20%

statistical significance in bivariate analysis into the Bayesian spatiotemporal negative binomial regression model. Variables included in the model were incidence risk by age group and village, age, bed net ownership, distance to health facilities, altitude, socio-economic status, area of residence and year of study. Spatial variation was included in the model as village specific random effects through latent observations of a spatial Gaussian process with mean zero and a covariance that assumed an exponential variation function of distance between two villages (Diggle et al., 1998). Temporal variation was modelled by a first order autoregressive process using monthly random effects. Bayesian models were fitted in OpenBugs version 3.1.2 (Imperial College and Medical Research Council London, UK) using Markov Chain Monte Carlo (MCMC) simulation. Covariate effects from the Bayesian geo-statistical model were considered statistically important when the credible intervals (CI) of the estimated regression coefficients did not include zero. Due to the nature of Bayesian statistical inference, we replace the terminology statistically significant by statistically important when reporting our results. Details of the Bayesian geostatistical temporal model are given in the Appendix 2.

3.2.5. Excess mortality attributed to slide positivity

To quantify the excess mortality rate (EMR), we used model coefficients for each age category to determine the mortality rates at different levels due to slide positivity rate (SPR). We calculated the difference between the mortality rates (MR) when SPR is greater than zero and when SPR is equal to zero.

$$EMR = MR(SPR > 0 - MR(SPR = 0)$$

The calculated EMR values were then plotted against SPR between 0.001 and 100 at intervals of 0.005.

3.3. Results

3.3.1. Descriptive statistics

During the period January 2007 to December 2012, there were a total of 375,447 uniquely registered residents in the study area contributing a total of 1,360,933 person years of observation (pyo). We observed an overall crude death rate of 13.8 deaths per 1,000 pyo, with a high of 18.8 deaths per 1,000 pyo in 2008 to a low of 10.2 deaths per 1,000 pyo in 2012. Malaria mortality was at 2.1 deaths per 1,000 pyo over the same period and followed a similar trend to the all-cause mortality rates. Alternatively slide positivity rate (SPR) during the same period was 52.5 % and followed a similar trend to the mortality rates over the years as shown in Table 3.1.

Table 3.1: All-cause mortality, malaria specific mortality and malaria risk rates by year

Year	Slide positivity rate (%)	Person years of observation*	All-cause death rate per 1,000 pyo (95% CI)	Malaria death rate per 1,000 pyo (95% CI)
2007	41.3 %	181537	15.5 (14.9, 16.1)	1.3 (1.2, 1.5)
2008	58.0 %	230374	18.8 (18.2, 19.3)	3.5 (3.3, 3.7)
2009	59.0 %	230373	15.6 (15.1, 16.2)	2.9 (2.6, 3.1)
2010	56.3 %	233871	12.4 (11.9, 12.8)	2.1 (1.9, 2.3)
2011	47.5 %	238524	10.9 (10.5, 11.3)	1.4 (1.2, 1.5)
2012	46.4 %	246254	10.2 (9.8, 10.6)	1.4 (1.2, 1.5)
Overall	**52.5 %**	**1360933**	**13.8 (13.6, 14.0)**	**2.1 (2.0, 1.3)**

*person years of observation (pyo) used for calculating all cause and malaria specific mortality only

Figure 3.2 and Figure 3.3 show the yearly and monthly trends in all-cause deaths, malaria specific deaths and malaria risk. The monthly trends show some peaks for all-cause and malaria specific deaths but not so distinct for SPR. The SPR rose sharply between December 2007 and January 2008 and stayed at these levels until the end of 2010, when there was a gradual drop from 2011 to 2012.

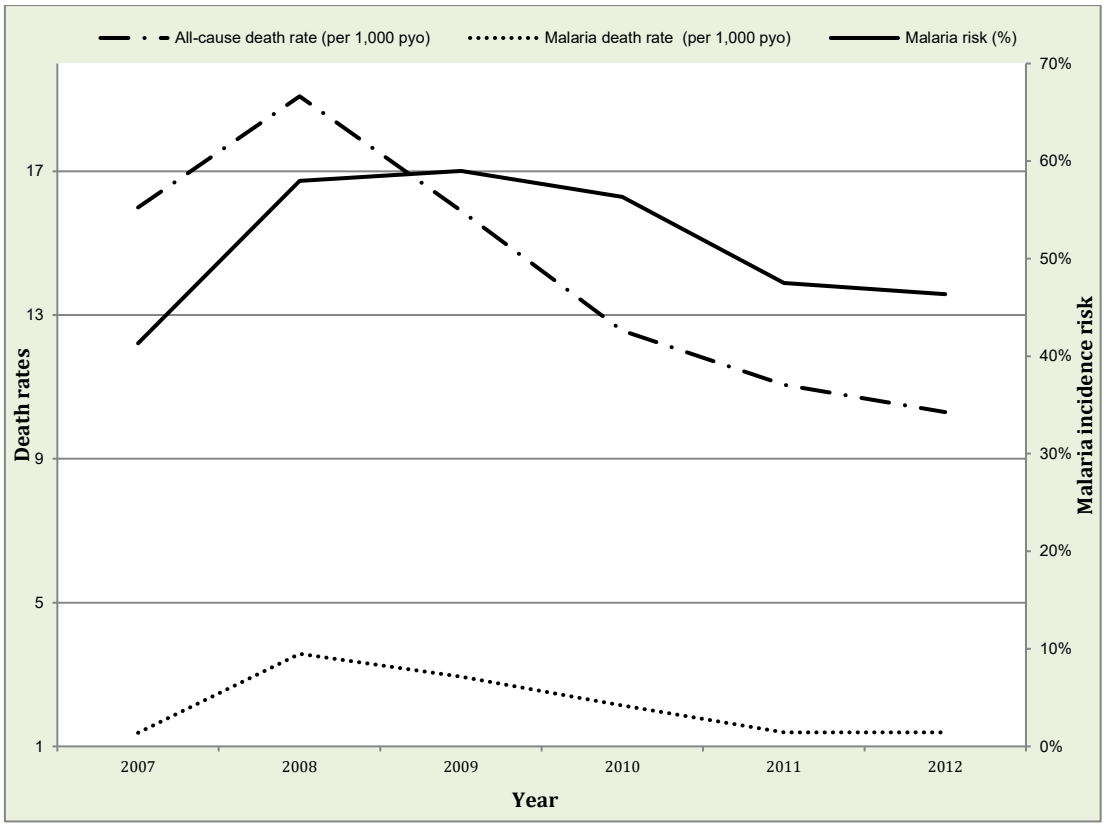


Figure 3.2: Yearly all-cause and malaria specific death rates versus malaria risk

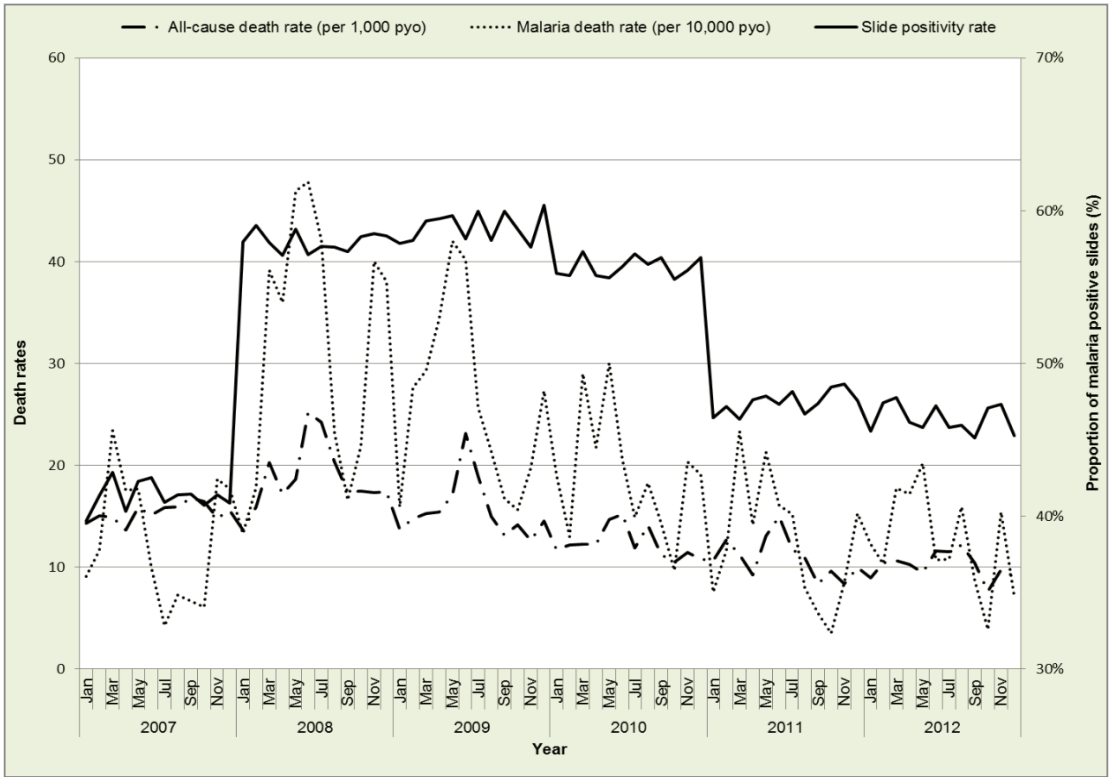


Figure 3.3: Monthly all-cause/malaria specific mortality versus slide positivity rate

Out of the total 18,729 deaths, 17,016 (91 %) had cause of death assigned by the InterVA. The top causes of deaths in the whole population were HIV/AIDS, malaria and pulmonary tuberculosis (Figure 3.4). Over the years, there was a downwards trend for most of the diseases as a proportion of overall deaths, there was however an upwards spike in malaria deaths in the years 2008/2009 which coincided with a drop for HIV/AIDS, acute respiratory infections, cardio-vascular illnesses and those grouped as others.

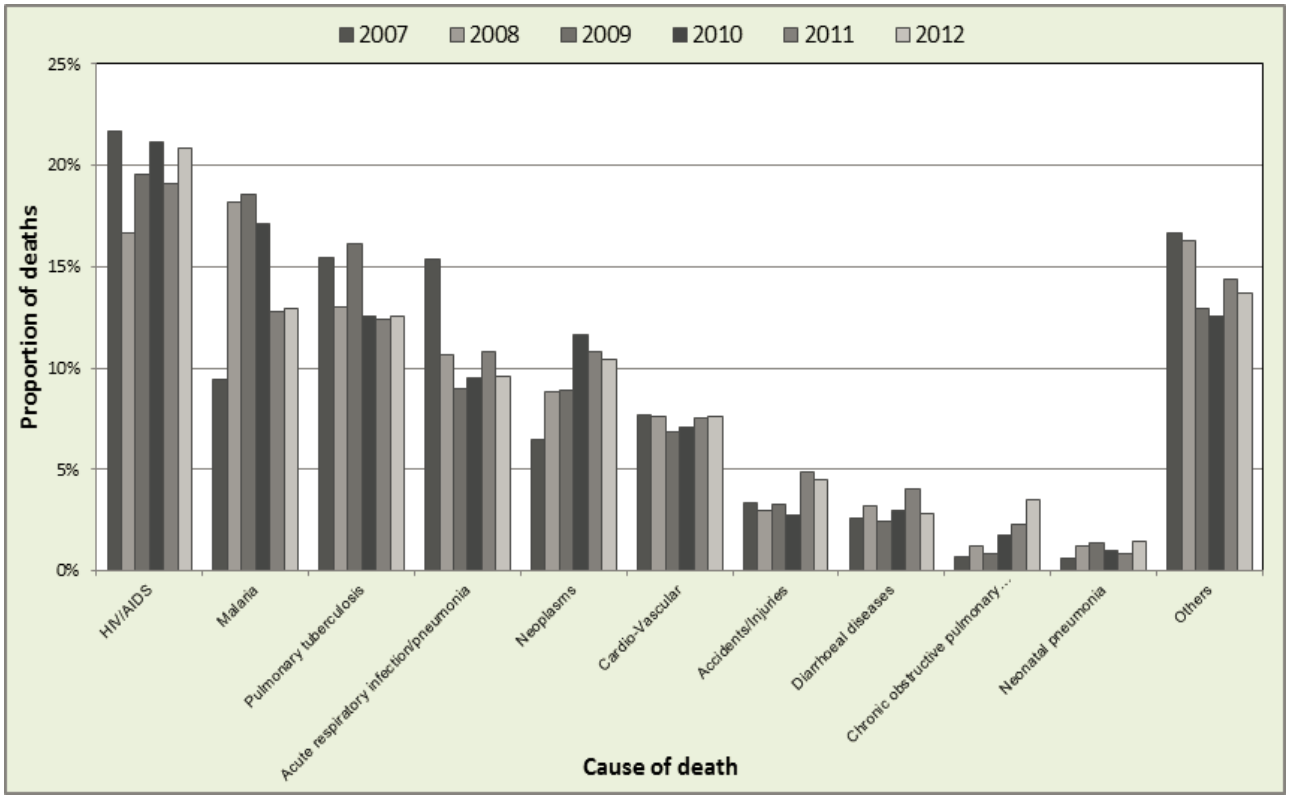


Figure 3.4: Main causes of death in the KHDSS

Figure 3.5 depicts the distribution of main causes of deaths including malaria among the different age groups. The top causes of mortality were pneumonia and births asphyxia for neonates, acute respiratory infection/pneumonia and malaria for infants, malaria and HIV/AIDS for both younger and older children, HIV/AIDS and pulmonary tuberculosis for adults, while for the elderly, the main causes were neoplasms and cardio-vascular diseases. Overall, the proportion of malaria was highest in the child age group (45 %), followed by older children (33 %) and infants (31 %).

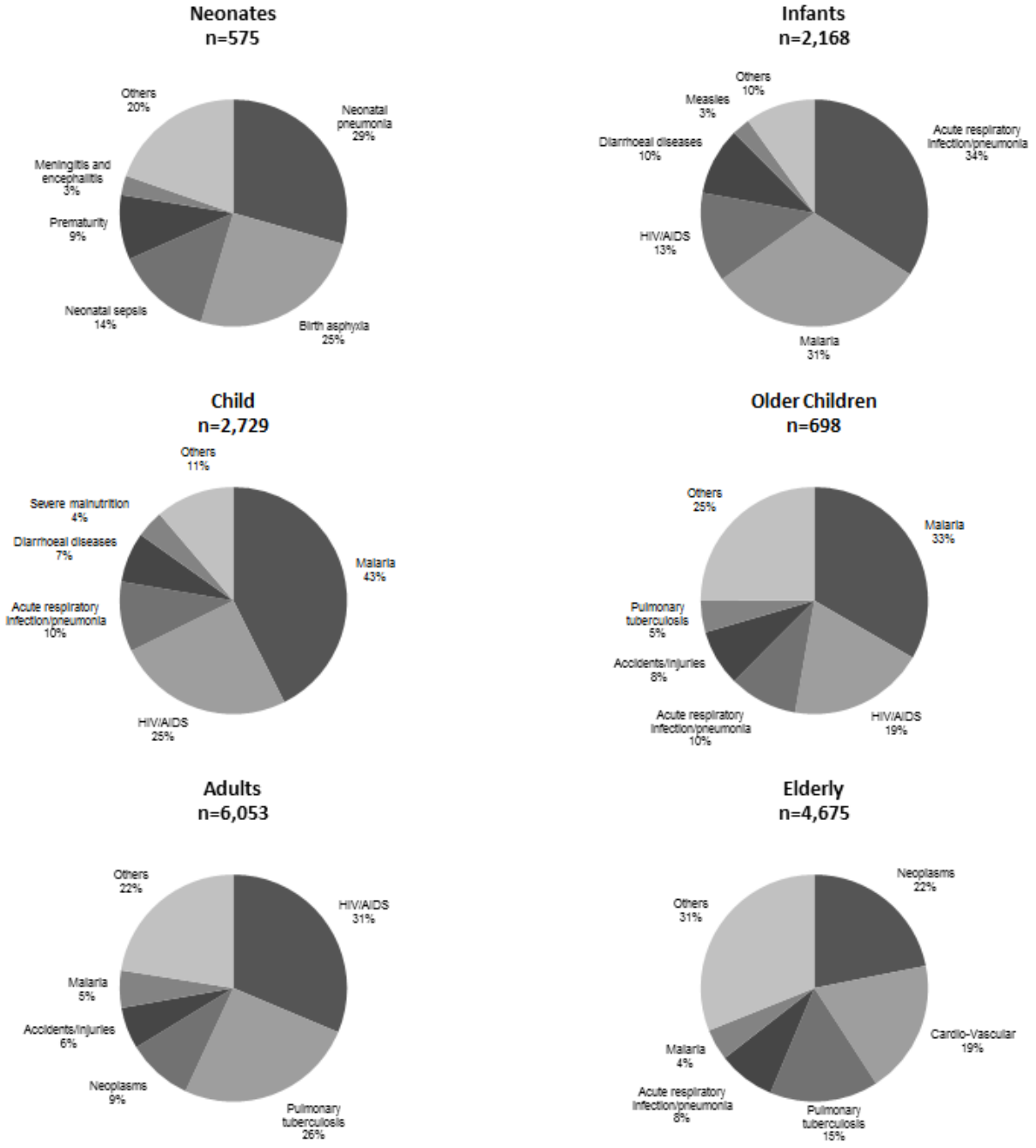


Figure 3.5: Main causes of death in the KHDSS by age groups

ITN ownership, distance to health facilities, socio-economic status, year of study, region and altitude were associated with either all-cause or malaria specific mortality in at least one of the age groups. For comparability purposes, these were included in all the seven Bayesian geostatistical models in both malaria-specific and all-cause mortality analysis.

3.3.2. Model based results

Table 3.2 shows the results from the Bayesian spatio-temporal negative binomial regression models for all-cause mortality by age groups at village levels. SPR was positively associated with all-cause mortality among infants, child, elderly and combined age group analysis. However, this association was important in the child (HR=4.29; 95 % CI: 2.78-13.29) and the combined age-group analysis (HR=1.55; 95 % CI: 1.04-2.80). SPR was not associated with all-cause mortality in neonates, older children and adults.

Average distance to health facilities was positively associated with increased all-cause mortality among all age groups. We observed higher mortality rates for distances greater than 1km or greater than 2km on average from health facilities. The association was important for neonates, infants, child and the combined analysis. Among the 3 levels of socio-economic status, there was reduced mortality risk in all age group with higher socio-economic status. The highest socio-economic status was associated with reduced all-cause mortality in neonates, children below 5 years and in the combined age group analysis. We observed a higher risk of mortality in the year 2008 compared to 2007 for all the ages which then reduced in the subsequent years and followed a downwards trend; however, in 2012 there was a slight reversal back to the pre-2008 levels among the neonates even though not statistically important. Higher altitude had a negative association with mortality in all the groups except the elderly and neonates but this association was only statistically important in the older children, adults and the combined age groups.

Estimated spatial variation at village level was higher than temporal variation in all age groups except the elderly. The minimum distance at which spatial variation was statistically important at 5 % ranged from 11.5 to 34.42.

Table 3.2: Posterior estimates (median) of the hazard ratio (HR) for predictors of all-cause mortality by age categories.

Covariates	Neonates (0-28days) HR (95% CI*)	Infants (1-11month) HR (95% CI)	Child (1-4yrs) HR (95% CI)	Older child (5-14yrs) HR (95% CI)	Adults (15-59yrs) HR (95% CI)	Elderly (60 +) HR (95% CI)	Overall** HR (95% CI)
Malaria risk	0.89 (0.13-5.70)	3.10 (0.36-13.12)	4.29 (2.78-13.29)	0.48 (0.15-2.05)	0.73 (0.39-1.42)	1.70 (0.79-4.45)	1.55 (1.04-2.80)
ITN	1.01 (1.00-1.02)	1.00 (0.99-1.01)	1.00 (0.99-1.003)	0.99 (0.98-1.01)	1.00 (0.997-1.002)	0.99 (0.99-1.00)	1.00 (0.998-1.001)
Distance to health facility							
0 – 1 km	1	1	1	1	1	1	1
1 – 2 km	1.21 (0.91-1.63)	1.24 (1.09-1.42)	1.12 (1.01-1.26)	1.01 (0.81-1.27)	1.15 (1.06-1.25)	1.00 (0.90-1.11)	1.11 (1.05-1.17)
>2 km	1.49 (1.11-2.11)	1.23 (1.07-1.43)	1.27 (1.12-1.44)	1.13 (0.93-1.45)	1.09 (0.99-1.20)	1.03 (0.92-1.16)	1.12 (1.05-1.20)
Ses							
Poorest	1	1	1	1	1	1	1
Poor	0.91 (0.70-1.17)	1.08 (0.95-1.23)	0.92 (0.83-1.03)	0.81 (0.64-1.01)	0.94 (0.86-1.02)	1.03 (0.94-1.13)	0.98 (0.93-1.03)
Least poor	0.64 (0.48-0.84)	0.97 (0.84-1.11)	0.79 (0.70-0.88)	0.83 (0.66-1.03)	0.94 (0.86-1.05)	0.99 (0.90-1.09)	0.94 (0.89-0.99)
Year							
2007	1	1	1	1	1	1	1
2008	1.25 (0.77-2.14)	1.29 (0.39-2.24)	1.23 (1.02-1.70)	1.31 (0.93-2.18)	1.05 (0.75-1.55)	1.07 (0.77-1.47)	1.15 (0.82-1.55)
2009	0.84 (0.49-1.47)	1.07 (0.28-1.85)	0.92 (0.69-1.19)	1.24 (0.73-2.13)	0.91 (0.64-1.33)	1.00 (0.72-1.44)	1.04 (0.67-1.48)
2010	0.54 (0.30-0.98)	0.84 (0.21-1.53)	0.74 (0.52-0.97)	1.00 (0.57-1.73)	0.73 (0.50-1.09)	0.92 (0.68-1.38)	0.88 (0.58-1.66)
2011	0.55 (0.31-0.97)	0.81 (0.25-1.81)	0.68 (0.47-0.92)	0.68 (0.38-1.12)	0.59 (0.41-0.89)	0.96 (0.70-1.44)	0.84 (0.57-2.01)
2012	1.02 (0.61-1.72)	0.72 (0.27-1.85)	0.62 (0.43-0.87)	0.85 (0.47-1.46)	0.56 (0.38-0.83)	0.89 (0.61-1.24)	0.80 (0.49-1.84)
Altitude							
1147 – 1243	1	1	1	1	1	1	1
1244 – 1293	1.41 (0.98-2.06)	1.07 (0.88-1.30)	1.07 (0.93-1.26)	0.95 (0.72-1.24)	0.97 (0.85-1.10)	0.96 (0.82-1.14)	0.97 (0.87-1.08)
1294 – 1327	1.33 (0.90-2.00)	0.85 (0.68-1.08)	0.94 (0.80-1.14)	0.74 (0.54-1.02)	0.84 (0.72-0.99)	0.91 (0.74-1.13)	0.85 (0.75-0.97)
1328 – 1365	1.46 (0.96-2.30)	0.79 (0.61-1.02)	0.99 (0.91-1.23)	0.68 (0.46-0.98)	0.92 (0.77-1.09)	1.01 (0.80-1.27)	0.88 (0.77-1.03)
>1365	1.43 (0.91-2.35)	0.75 (0.55-1.01)	0.89 (0.75-1.14)	0.65 (0.42-0.94)	0.88 (0.72-1.07)	1.01 (0.78-1.31)	0.87 (0.74-1.03)
Spatial Variation	0.44 (0.14-1.90)	2.36 (0.70-8.93)	2.04 (0.72-7.61)	0.74 (0.22-3.59)	0.99 (0.33-4.42)	0.52 (0.20-1.88)	1.69 (0.76-6.15)
Temporal Variation	0.14 (0.08-0.25)	0.10 (0.07-0.17)	0.11 (0.07-0.17)	0.13 (0.08-0.22)	0.06 (0.04-0.08)	0.07 (0.04-0.10)	0.05 (0.04-0.08)
Range***	34.41 (8.77-95.37)	11.15 (8.10-39.18)	11.38 (8.11-40.52)	17.19 (8.27-84.45)	13.18 (8.16-55.67)	20.79 (8.45-83.93)	11.40 (8.12-40.42)

* CI=credible interval

** In addition to the variables above we also adjusted for age in the overall model

*** Minimum distance in kilometres at which spatial variation is statistically important at 5%

In modelling malaria-specific mortality in relation to SPR, we developed a Bayesian spatio-temporal model with similar variables used in the all-cause mortality model as shown in Table 3.3. Neonates were excluded from the malaria-specific modelling because there were no deaths attributed to malaria by InterVA in this group. Malaria mortality increased with an increase in SPR among infants, child and overall analysis. SPR was strongly associated with increased risk of malaria mortality among the child group (HR=9.48; 96 % CI: 5.11-37.94); however, in, it was strongly associated with a reduction in malaria specific mortality (HR=0.02; 95 % CI: 0.003-0.33).

Similar to the all-cause mortality above, reduced average distance to health facilities, higher socio-economic status and year of study were associated with reduced risk of malaria mortality. In all models, spatial variation was higher than temporal variation with the exception of the model corresponding to the 60+ age group. The minimum distance at which spatial variation for malaria mortality by village was statistically important at 5 %, ranged from 10.74km in the child age group to a high of 36km among the elderly.

Table 3.3: Posterior estimates (median) of the hazard ratio (HR) for predictors of cause-specific mortality by age categories.

Covariate	Infants (1-11month) HR (95% CI*)	Child (1-4yrs) HR (95% CI)	Older child (5-14yrs) HR (95% CI)	Adults (15-59yrs) HR (95% CI)	Elderly (60 +) HR (95% CI)	Overall** HR (95% CI)
Malaria risk	1.36 (0.23-9.85)	9.48 (5.11-37.94)	0.02 (0.003-0.33)	0.27 (0.02-3.24)	0.59 (0.01-13.15)	1.37 (0.51-3.73)
ITN	1.01 (1.00-1.02)	1.0 (0.99-1.01)	1.01 (0.99-1.03)	0.99 (0.98-1.01)	1.02 (1.00-1.04)	1.00 (0.99-1.01)
Distance to health facility						
0 – 1 km	1	1	1	1	1	1
1 – 2 km	1.19 (0.89-1.60)	1.26 (1.05-1.45)	1.10 (0.76-2.04)	1.02 (0.72-1.45)	0.98 (0.50-2.01)	1.12 (0.95-1.30)
>2 km	1.18 (0.85-1.64)	1.42 (1.35-1.66)	0.93 (0.75-1.20)	0.94 (0.67-1.10)	1.24 (0.63-2.51)	1.14 (0.95-1.37)
Ses						
Poorest	1	1	1	1	1	1
Poor	1.03 (0.79-1.37)	0.94 (0.77-1.08)	1.01 (0.55-1.46)	1.03 (0.66-1.58)	1.10 (0.60-2.13)	0.99 (0.86-1.15)
Least poor	0.94 (0.70-1.26)	0.76 (0.62-0.89)	0.58 (0.41-1.07)	1.27 (0.82-1.95)	0.87 (0.53-1.86)	0.92 (0.79-1.07)
Year						
2007	1	1	1	1	1	1
2008	9.32 (3.34-57.79)	2.34 (1.36-4.07)	1.26 (0.63-4.03)	1.60 (0.79-3.36)	1.73 (0.49-2.67)	2.12 (0.93-4.14)
2009	7.21 (2.56-47.57)	1.57 (1.08-2.51)	4.59 (1.81-7.23)	2.13 (1.52-4.36)	3.28 (1.07-58.62)	1.86 (0.57-3.89)
2010	6.95 (2.36-45.32)	1.87 (1.18-2.70)	5.19 (1.85-15.48)	1.43 (0.70-2.54)	5.75 (1.66-15.5)	2.02 (0.55-4.02)
2011	6.07 (1.99-37.2)	1.74 (1.25-2.87)	0.77 (0.46-2.48)	0.79 (0.32-1.71)	3.00 (1.16-30.4)	1.45 (0.32-3.02)
2012	5.31 (1.66-36.72)	2.09 (1.33-3.19)	2.09 (1.01-6.14)	0.95 (0.43-2.05)	2.68 (0.99-46.27)	1.60 (0.34-3.28)
Altitude						
1147 – 1243	1	1	1	1	1	1
1244 – 1293	0.82 (0.54-1.24)	0.69 (0.56-0.92)	0.73 (0.39-1.62)	0.61 (0.28-1.14)	1.46 (0.93-5.25)	0.70 (0.54-0.93)
1294 – 1327	0.63 (0.40-1.03)	0.72 (0.57-0.98)	1.14 (0.55-2.43)	0.71 (0.39-1.31)	1.04 (0.28-3.73)	0.68 (0.50-0.92)
1328 – 1365	0.58 (0.35-0.99)	0.71 (0.56-1.01)	0.65 (0.42-1.64)	0.73 (0.37-1.42)	1.63 (0.55-5.65)	0.64 (0.45-0.88)
>1365	0.39 (0.22-0.76)	0.64 (0.46-0.91)	1.08 (0.50-2.67)	0.57 (0.27-1.15)	2.74 (1.14-10.81)	0.58 (0.40-0.85)
Spatial Variation	5.87 (1.00-29.11)	5.66 (1.30-21.34)	0.73 (0.22-5.14)	0.46 (0.17-2.91)	0.40 (0.15-1.81)	8.83 (4.11-29.28)
Temporal Variation	0.15 (0.09-0.28)	0.18 (0.11-1.30)	0.41 (0.16-1.28)	0.29 (0.14-0.61)	0.42 (0.17-1.38)	0.17 (0.11-0.27)
Range***	10.95 (8.09-40.52)	10.74 (8.08-36.21)	18.37 (8.27-82.37)	30.11 (8.56-93.64)	36.0 (8.83-95.26)	10.9 (8.09-34.21)

* CI=credible interval

** In addition to the variables above we also adjusted for age in the overall model

*** Minimum distance in kilometres at which spatial variation is statistically important at 5%

On plotting excess mortality as a function of malaria transmission (SPR), we showed that infants, children 1-4 years, elderly adults and the combined population exhibited an increase in mortality rate with SPR. However a protective effect was noted among children 4-14 years and adults 15-59 years. The highest burden of mortality attributable to SPR was noted among children 1-4 years followed by infants 1-11 months (Figure 3.6).

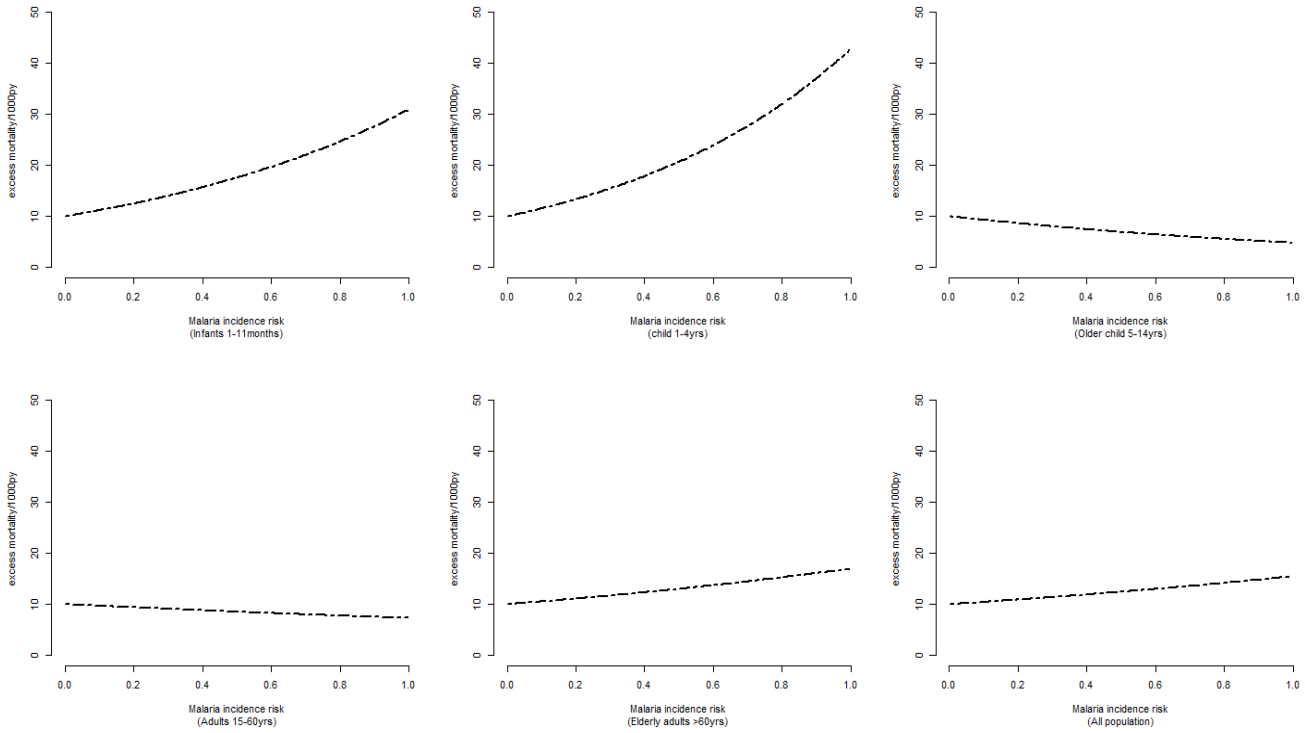


Figure 3.6: Age specific patterns of excess mortality due to malaria incidence

3.4. Discussion

Several studies have investigated the effect of malaria transmission on mortality using mostly entomologic inoculation rate or prevalence data as a measure of transmission (Amek, 2013; Gething et al., 2016; Rumisha et al., 2014). Entomological data is quite sparse while prevalence data does not reflect seasonal variations of transmission. Our study is the first to link HDSS malaria mortality and health facility incidence data collected continuously in a well-defined geographical area. Using these datasets well aligned in space and time, we investigated SPR as

a measure of transmission in relation to all-cause and malaria-specific mortality in the KHDSS area of western Kenya employing rigorous Bayesian spatio-temporal models to account for variation in space and time. We adjusted for person time observed as discrete monthly intervals, socio-economic status, ITN ownership, average distance to health facilities, altitude and year of study. In estimating excess mortality due to malaria transmission, it was noted that small changes in slide positivity rate (SPR), results in significant increases in overall mortality in this population.

In modelling the relationship between SPR and mortality, we found that SPR predicts all-cause mortality in the whole population. Among different age-groups, SPR was an important measure of both all-cause and malaria-specific mortality for children aged 1-4 years implying that malaria infection in under 5 children contributes greatly to overall mortality. Other studies have also shown a positive association between malaria transmission and all-cause mortality in this age group (Amek, 2013; Gething et al., 2016; Ross & Smith, 2006; Rumisha et al., 2014), however, were focused mostly on the relationship between all-cause mortality and EIR or prevalence. Children aged 1-4 years are most affected by malaria mortality due to lower immunity compared to neonates or children below 6 months (Malhotra et al., 2009) (Malhotra et al. 2009). The effect of SPR on malaria-mortality among children aged 1-4 years could also be a result of delayed access to treatment and the magnitude of malaria burden in this group. Decreasing malaria incidence through proper and timely treatment and management can therefore greatly reduce both all-cause and malaria specific mortality. Slide positivity rate as a measure of malaria exposure appears to be a protective factor for malaria-specific mortality in older children (5-14 years), most likely due to acquired immunity at lower ages (Bejon et al., 2009; Stanisic et al., 2014). This explanation is supported by data from the same area indicating that *Plasmodium falciparum* parasite prevalence is highest in this group (Khagayi et al., 2019). These school going children are less mobile compared to younger children who move with their working parents (Adazu et al., 2005), and therefore the constant exposure to infection increases

their immunity. It is important to note that even though there was low overall mortality in older children; malaria deaths as a proportion of all deaths in this group were still high and increased over the years. Desai et al also reported that malaria deaths in the 5-14 age group increased over the years; an indication reflecting either behavioural factors or reduced attention towards this age-group from control programs that target mainly pregnant women and children under five years (Desai et al., 2014) (Desai et al. 2014).

Our study supports earlier findings in the same study area where EIR as a measure of transmission was seen to drive mortality in children under 5 years (Amek et al., 2018) but not in the 5-14 age group. Of note, is that in our study, SPR (OR=4.29 and OR=9.48) shows a stronger effect than EIR (OR=1.58 and OR=1.97) for both all cause and malaria specific mortality respectively. With proper diagnosis and data collection, SPR can be collected more easily using fewer resources at local levels than EIR. This highlights the importance of incidence as a driver of malaria transmission and consequently a better measure of exposure that can be used to monitor progress towards malaria control in areas of high endemicity.

During the study period, there was an upward spike in all-cause mortality, malaria specific mortality and SPR in 2008, followed by a steady decline between 2009 and 2012. The mortality increase in the year 2008 was observed in previous studies carried out in the same area and attributed to the effect of post-election violence in Kenya at the end of 2007 (Desai et al., 2014; Feikin et al., 2010; Hamel et al., 2011). The rise in all-cause/malaria-specific mortality and SPR could be attributed to these clashes, which disrupted health services, interfered with the supply and provision of antimalarial drugs and led to a surge in malaria among other infectious diseases. Our findings echo what has been observed in other conflict areas in Africa and demonstrated the impact of conflicts on malaria burden (Sedda et al., 2015). Malaria as a proportion of all deaths was most affected since its increase in 2008-2009 coincided with a drop in HIV/AIDS, acute respiratory infections and cardio-vascular diseases among others showing that malaria can easily bounce back more forcefully whenever control efforts are interrupted.

Malaria-specific mortality reduction in the other age groups can be attributed to the scale up of interventions, increased coverage of ITNs, prompt and improved malaria treatment using artemisinin-based combination therapies (ACT) and indoor residual spraying (Bhatt et al., 2015; Division of Malaria Control [Ministry of Public Health and Sanitation] et al., 2011; Hamel et al., 2011). All-cause mortality reduction could be attributed to a reduction in malaria-specific mortality, scaled up provision of antiretroviral therapy (ART) and improvements in health service delivery that saw declines in infectious diseases and HIV/AIDS (Desai et al., 2014; Gargano et al., 2012). The reduction in mortality is expected to improve in the future since many health policy decisions, care and management have been devolved to the county levels with a view of tailoring solutions to suit local needs (Masaba et al., 2020).

ITN ownership was not related to SPR. The expected individual effect of ITNs on mortality could have been lost due to aggregating the data at village level. Distance to health facilities was an important factor in determining all-cause mortality across all age groups; showing that relatively small differences in distance can substantially affect mortality, especially in younger children. A study in the KHDSS on paediatric hospitalization argued that longer distances act as barriers to seeking care making people stay away even if they are sick (Feikin et al., 2009). Access to antenatal services, which is an important factor for neonatal survival is largely influenced by distance to health facilities (Karra et al., 2016) and may explain the large effect among neonates. Comparable to our rural setting, other studies have found similar results in sub-Saharan Africa (Kadobera et al., 2012; Ombok et al., 2010; Rutherford, 2009; Schoeps et al., 2011).

People at lower socio-economic status experienced relatively higher mortality compared to those at a higher status. The effect was important among neonates for all-cause mortality and in children under the age of five for both all-cause and malaria specific mortality. This is consistent with other studies done in similar HDSS settings (INDEPTH Network, 2005). The poorest would most likely live in houses that are not well constructed hence offering little

protection against vectors and are less likely to pay for effective treatment. This vulnerability is more pronounced in younger children even with small scale economic differences visible through simple asset ownership (Sachs & Malaney, 2002).

Altitude's negative association with both all-cause and malaria-specific mortality supports findings showing that malaria vector abundance reduces with altitude (Githeko et al., 2006). Previous studies in the western Kenya region indicated that middle level and low altitude areas experienced higher mortality rates compared to the higher regions (Ombok et al., 2010).

A strong spatial variation was estimated in the malaria specific model for infants, children u 1-4 years and overall model, suggesting that mortality is influenced by spatially structured exposures. This finding has also been noted in previous studies in the same region and other places in sub-Saharan Africa (Amek et al., 2018; Kasasa et al., 2013; Rumisha et al., 2014). We also noted that for all age groups, spatial variation was higher than temporal variation signifying a reduced seasonal influence on mortality compared to spatial variability.

The use of verbal autopsy as a tool for determining cause of death has been criticized as not being very specific and can either under-estimate or over-estimate malaria (Murray et al., 2012). The InterVA tool for verbal autopsy has however undergone rigorous tests and improvements recently to take care of physician failings, differences in high and low malaria transmission areas and found to be in high concordance with physician coding (Byass et al., 2015). It also agrees with other population-based projections in determining malaria as a cause of deaths, hence a useful tool for ascertaining cause of death in low-income settings that lack proper vital registration systems. In our study, the SPR estimates may be biased due to missing patients who do not make it to health facilities, more so the older ages with reported lower health seeking behaviour for fever related illnesses relative to younger children (Bigogo et al., 2010). At the same time, microscopy case confirmation has been shown to have lower sensitivity (Lo et al., 2015) and the SPR may be influenced by changes in other febrile illnesses by inflating the denominator (Boyce et al., 2016). Using SPR from children below the age of 14 years as an

indicator for transmission in the whole population could result in a bias since the behavioural and biological characteristics of children may not match exactly those of the whole population. However, we anticipated that the SPR of adults would be correlated at village level with that of children under the age of 14 as they come from the same community as the adults from whom we infer association between under 14 children's SPR and adult mortality due to the strong environmental, climatic and other spatially explicit factors that drive malaria transmission (Mboera et al., 2010). The sudden rise of SPR although not due to changes in the diagnostic methods could have been due to unobserved effects due to changes in health systems and may have biased the results to reflect a stronger effect in the association with mortality. By using longitudinal data over a long period of time, consistent methodology, and rigorous quality control for microscopy diagnosis and assuming that SPR represents the level of malaria exposure in the population, we offset some of the shortcoming of using sentinel health facilities data.

3.5. **Conclusion**

Our study showed that slide positivity rate is significantly associated with all-cause/malaria-specific mortality in this region of western Kenya. By quantifying excess mortality due to malaria, we show that small changes in malaria incidence can substantially reduce deaths.

As the fight towards malaria control and elimination continues, incidence risk data from sentinel health facilities can be used as a measure of exposure to assess, monitor and quantify both all-cause and malaria-specific mortality in low resource settings.

Declarations

Authors' contributions

SK conceived the study, analysed the data and drafted the initial manuscript; PV conceived the study and supported the data analysis. All authors gave intellectual inputs in analysis, revised the drafts, read and approved the final manuscript.

Acknowledgements

We would like to thank the entire KHDSS community for participating in the longitudinal surveillance, providing this information through the many years, and supporting our field activities; the entire KHDSS staff, past and present for assisting in managing the platform, data collection/management and archiving; the staff and management of all the health facilities where we collect malaria incidence data

Ethical statement

The study protocol for the KHDSS has been approved by the Kenya Medical Research Institute (#1801) and the Centers for Disease Control and Prevention (#3308) institutional review boards. Written informed consent was obtained from the heads of compounds in conducting the household interviews while hospital data was only collected from individuals who also gave an additional written informed consent or from the caretaker in-charge if the participants were under 18 years of age.

Data Availability

Data were obtained with permission of the KHDSS Steering committee. The data are available from the Kenya Medical Research Institutes' Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data. Data are from the HDSS study whose authors may be contacted at Gbigogo@kemricdc.org or Skhagayi@kemricdc.org. This article is published with the permission of the director Kenya Medical Research Institute. The KHDSS is a member of the INDEPTH Network.

3.6. Appendix

3.6.1. Spatial and temporal description of data used in the study

Data type	Description	Source	Temporal resolution
Malaria slide positivity rates	Measured as the number of malaria positive slides out of the total blood slides examined by light microscopy.	Sentinel health facilities	Continuously (Jan 2007-Dec2012)
Cause of death (Malaria-specific)	Actual deaths are reported continuously by village reporters	Household level	Continuously (Jan 2007-Dec2012)
Socioeconomic status	Constructed based on household asset ownership using a composite score, the multiple correspondence analysis (MCA) technique and categorized into 3 quintiles as least poor for the well off, poor for the average and poorest for the lowest rank	Household level	Yearly (2007-2012)
Person time	The HDSS collects data on an initial population at the start of the observation period followed by subsequent 4 monthly surveillance cycles that provide data on births, deaths, in-migration and out-migrations. Using these, person time of observation in years was calculated as the total time spent by a registered KHDSS resident in the study area during the study period	Household level	Continuously (Jan 2007-Dec2012)
Bed net ownership	Calculated as the percentage of households per village owning at least one net for every two people	Household level	Yearly (2007-2012)
Altitude	Extracted from the Shuttle Radar Topography Mission (SRTM) by the U.S. Geological Survey - Earth Resources Observation and Science (USGS EROS) Data Center (https://eros.usgs.gov/elevation-products)	Household level	Once
Distance to health facility	Calculated as the Euclidean difference between the household and the nearest health facility in kilometres, aggregated at village level and classified as less than 1km, 1 to 2km and greater than 2km	Household level	Once
All-cause mortality	From the KHDSS continuous 4 monthly cycles, data on all deaths among the registered residents is collected to provide the number of deaths in the population and by age groups.	Household level	Continuously (Jan 2007-Dec2012)

3.6.2. Bayesian model formulation

Let $\mathbf{Y}_{jt}$ be the average number of deaths (all-cause or malaria-specific) in village *j* at time interval *t*. We assume that $\mathbf{Y}_{jt}$ arises from a negative binomial distribution.

$$\mathbf{Y}_{jt} \sim \text{dnegbin}(\mathbf{P}_{jt}, \boldsymbol{r})$$

Where P_{jt}, is the proportion of deaths occurring in village *j* at time interval *t* and *r* is the dispersion parameter with,

$$\mu_j = r\frac{1-p}{p} \text{ and } \sigma_i^2 = r(1-p)p^{-2}.$$

We modelled the association above between covariates (*X*) and mortality status of individuals by village on the logit, as

$$logit(\mu_{jt}) = logit(N_{jt}) + \beta_0 + \sum_1^k \beta_u X_u + \phi_j + \varepsilon_t \qquad , u=1,2,\ldots\ldots k$$

where μ_{jt} is the number of deaths in each village at time *t*, N_{jt} the total person time contributed by persons in each village as discrete months, β_i the regression coefficients, ϕ_j the village specific spatial effects and ε_t the temporal (monthly) random effects.

We assumed that ϕ_j are parameters from a latent spatial process modelled by a Gaussian distribution with covariance matrix quantifying the relation between any pair of villages as a function of their distance irrespective of the direction using an exponential correlation function, that is $\emptyset \sim \text{MVN}(0, \sum)$, $\sum_{kl} = \sigma_1^2 \exp(-\rho d_{kl})$ where σ_1^2 is the spatial variation, d_{kl} is the distance between villages *k* and *l*, and ρ is the rate of correlation decay with increasing distance. The minimum distance at which the spatial variation is less than 5% is called range and can be obtained from the value $3/\rho$ (Ecker & Gelfand, 1997). Temporal effect (ε_t) was modeled by an autoregressive process of order 2. We specified non-informative normal prior distributions with mean zero and large variance for the β_i *i*=1,…… regression coefficients, an inverse gamma prior for $r \sim IG(1.01, 0.001)$, an inverse gamma priors for σ_e^2 and σ^2. A gamma prior for ρ, that is $\sigma_e^2, \sigma^2 \sim IG(2.01, 1.01)$ and $\rho \sim G(0.01, 0.01)$.

The model was fitted using Markov Chain Monte Carlo (MCMC) simulation algorithm in OpenBugs version 3.1.2 (Imperial College and Medical Council, London, UK) to estimate model parameters (Gelfand & Smith, 1990). We ran a single chain sampler discarding the first 10,000 iterations. Convergence was assessed by Gelman-Rubin diagnostic (Gelman & Rubin, 1992) and attained at 100,000 iterations.

4. Chapter 4: Modelling the relationship between malaria prevalence as a measure of transmission and mortality across age groups.

Sammy Khagayi[1,2,3], Meghna Desai[4], Nyaguara Amek[1], Vincent Were[1], Eric D. Onyango[1], Christopher Odero[1], Kephas Otieno[1], Godfrey Bigogo[1], Simon Kariuki[1] Aaron Samuels[4,5], Stephen Munga[1], Frank Odhiambo[1], Mary Hamel[4], Laurence Slutsker[4], John Gimnig[4], and Penelope Vounatsou[2,3§]

1. Kenya Medical Research Institute-Center for Global Health Research, Kisumu, Kenya
2. Swiss Tropical and Public Health Institute, Basel, Switzerland
3. University of Basel, Basel, Switzerland
4. Centers for Disease Control and Prevention, Center for Global Health, Division of Parasitic Diseases and Malaria, Malaria branch, Atlanta, GA, USA
5. Centers for Disease Control and Prevention, Kisumu, Kenya

[§]Corresponding author

This manuscript has been published in *Malaria Journal* August 2019

Abstract

Background

Parasite prevalence has been used widely as a measure of malaria transmission, especially in malaria endemic areas. However, its contribution and relationship to malaria mortality across different age groups has not been well investigated. Previous studies in a health and demographic surveillance systems (HDSS) platform in western Kenya quantified the contribution of incidence and entomological inoculation rates (EIR) to mortality. The study assessed the relationship between outcomes of malaria parasitaemia surveys and mortality across age groups.

Methods

Parasitological data from annual cross-sectional surveys from the Kisumu HDSS between 2007 and 2015 were used to determine malaria parasite prevalence (PP) and clinical malaria (parasites plus reported fever within 24 hours or temperature above 37.50C). Household surveys and verbal autopsy (VA) were used to obtain data on all-cause and malaria-specific mortality. Bayesian negative binomial geo-statistical regression models were used to investigate the association of PP/clinical malaria with mortality across different age groups. Estimates based on yearly data were compared with those from aggregated data over four to five-year periods, which is the typical period that mortality data are available from national demographic and health surveys.

Results

Using 5-year aggregated data we established associations between parasite prevalence and malaria-specific mortality in the whole population ($RR_{malaria}$=1.66; 95 % Bayesian Credible Intervals: 1.07-2.54) and children 1-4 years ($RR_{malaria}$=2.29; 1.17-4.29). While clinical malaria was associated with both all-cause and malaria-specific mortality in combined ages ($RR_{all-cause}$=1.32; 1.01-1.74); ($RR_{malaria}$=2.50; 1.27-4.81), children 1-4 years ($RR_{all-cause}$=1.89; 1.00-3.51); ($RR_{malaria}$=3.37; 1.23-8.93) and in older children 5-14 years ($RR_{all-cause}$=3.94; 1.34-11.10); ($RR_{malaria}$=7.56; 1.20-39.54), no association was found among neonates, adults (15-59 years) and the elderly (60+ years). Distance to health facilities, socioeconomic status, elevation and survey year were important factors for all-cause and malaria-specific mortality.

Conclusion

Malaria parasitaemia from cross-sectional surveys was associated with mortality across age groups over four-to-five-year periods with clinical malaria more strongly associated with mortality than parasite prevalence. This effect was stronger in children 5-14 years compared to other age-groups. Further analyses of data from other HDSS sites or similar platforms would be useful in investigating the relationship between malaria and mortality across different endemicity levels.

Keywords: Malaria, mortality, parasite prevalence, Bayesian spatio-temporal, health and demographic surveillance system**.**

4.1. Introduction

There has been a substantial reduction in malaria related mortality worldwide over the last decade, however, the burden is still disproportionately felt in sub-Saharan Africa (SSA) (Bhatt et al., 2015). Due to the high burden in children and pregnant women (World Health Organization, 2016), malaria control intervention resources in previous years have been targeted to these vulnerable populations. Increased quality data on malaria infection dynamics and mortality across all ages has created an increased awareness of the burden of disease amongst the other populations, and policies have been expanded to ensure universal coverage with effective vector control methods (e.g. long-lasting insecticidal nets [LLIN]), availability of diagnostics (e.g. rapid diagnostic tests [RDT]), and availability of appropriate treatments (e.g. artemisinin-based combination therapy [ACTs]) to all (Sankoh & Binka, 2005).

There is evidence that the malaria burden in older children and adults in terms of mortality and parasite prevalence (Desai et al., 2014; Okiro et al., 2009; Walldorf et al., 2015) is higher than had been thought of previously. With data analysed from a health and demographic surveillance system (HDSS) in western Kenya run by the Kenya Medical Research Institute (KEMRI) and Centers for Disease Control and Prevention (CDC) showing that, largely due to increased malaria/HIV prevention and treatment interventions, malaria mortality rates decreased in young children and persons aged ≥15 years, but remained stable in 5-14 year olds (Desai et al., 2014); suggesting that malaria control efforts should be intensified in this group. Furthermore, older children and adults have been shown to act as reservoirs of transmission due to high levels of asymptomatic infections (Zhou et al., 2016), supporting the current policy of universal coverage of malaria control interventions.

Measuring malaria transmission intensity and its effect on mortality can be used to monitor disease burden and assess the impact of interventions and control programs. This has been done previously using entomological inoculation rates (EIR) (Amek et al., 2018; Rumisha et al.,

2014); however, measuring EIR is expensive, time consuming and is often imprecise, particularly in low transmission settings. Other measures of malaria transmission that have been explored before in relation to mortality include slide positivity rate (SPR), parasite prevalence, disease incidence, sporozoite rate, and vectoral capacity (Khagayi et al., 2017; O'Meara et al., 2008; Ross & Smith, 2006; Smith et al., 2001).

Malaria parasite prevalence (PP) surveys carried out mostly during peak transmission times through representative sampling of populations are a preferred method for measuring malaria burden because reporting from weak or non-existent health systems is inadequate to measure incidence (Corsi et al., 2012), at the same time health facilities do not capture asymptomatic infections which are important for malaria transmission. Furthermore PP survey data are easier to interpret and less prone to uncertainty compared to other measures (Snow, 2014). These surveys are however limited in their ability to capture malaria morbidity, seasonality of transmission and monitor temporal trends from surveys that are not seasonally aligned (Moss et al., 2015). With regular, consistent survey intervals, stringent methodology in sampling and diagnosis, PP surveys can provide measures of malaria transmission which are useful to policy makers.

Due to their nature, HDSS sites can be used to collect data that are well aligned in space and time so as to investigate variations in malaria transmission in relation to morbidity and mortality. They provide data on mortality across age groups, and in conjunction with PP surveys offer a unique avenue through which we can investigate and understand the relationship between malaria transmission and mortality while taking into consideration spatio-temporal factors (Moss et al., 2015; Streatfield et al., 2014; Ye et al., 2012), and hence monitor the impact of interventions over time. One such project, the Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) investigated the effect of EIR as a measure of exposure and its effect on mortality in several HDSS sites in Africa and showed that small changes in

transmission dynamics as measured by EIR, impact greatly on mortality (Amek, 2013; Amek et al., 2018; Rumisha et al., 2014).

In this study, we sought to understand how malaria parasite prevalence and clinical malaria translate into mortality and consequently help inform national control programs on how to best use their survey data in estimating mortality. We investigated the relation between malaria prevalence and mortality across all age-groups using Bayesian geostatistical models on data collected between 2007 and 2015 from the KEMRI and CDC HDSS (KHDSS) site in western Kenya. To the best of our knowledge, there have been no studies done to investigate the usefulness of parasite prevalence as a proxy for malaria transmission and its association with mortality using data that are well aligned in space and time across different age groups in similar settings.

4.2. Methods and data

4.2.1. Study area and population

The KHDSS located in Siaya County of western Kenya follows a population of over 240,000 people as of mid-2015 in an area of over 700km^2 (Odhiambo et al., 2012). This HDSS is located in a malaria endemic zone with a high burden of HIV/AIDS compared to the rest of the country (Division of Malaria Control [Ministry of Public Health and Sanitation] et al., 2011; National AIDS and STI Control Programme (NASCOP), 2014).

From the HDSS, we collected data on an initial population at baseline, which was followed by subsequent 4 monthly cycles every year during which data were collected on births, deaths, in-migration and out-migrations. These data were used to estimate person-years of observation (pyo) that served as a denominator to calculate mortality rates. Verbal autopsy (VA) was used to determine malaria-specific mortality rates. The methods used for verbal autopsy have been described in detail elsewhere(Adazu et al., 2005; Odhiambo et al., 2012); it involves capturing

data on a deceased person's last illness, signs, symptoms and medical history which is then used to determine the most probable cause of death using a computer-based Bayesian expert algorithm called InterVA (Byass et al., 2012).

4.2.2. Malaria prevalence

Annual all-age malaria and anaemia prevalence surveys were conducted by randomly sampling compounds within the HDSS, and testing all consenting members of the compound for malaria by blood smear microscopy, from the population during the peak malaria transmission period in July. Details of the sampling by year are shown in Additional file 1. Trained interviewers then visited the compounds, administered a questionnaire to collect information on demographics, risk factors for malaria infection, healthcare seeking, previous illness, socioeconomic status, LLIN ownership/use, and collected a blood sample to prepare thick and thin smears for microscopy. The blood slides were transported to a central laboratory, stained with 10% Giemsa and examined for malaria parasites by expert microscopists.

We considered two measures of transmission; prevalence of malaria parasites and clinical malaria for comparative purposes. Parasite prevalence by age group, village, and year was defined as the proportion of participants in each village that had malaria by microscopy out of all the participants from the same village who were tested for malaria. Similarly, clinical malaria prevalence was defined as the proportion of participants in each village who had malaria parasites of any density by microscopy in combination with either a reported fever in the previous 24 hours or a temperature of 37.5^0C and above out of all those tested.

4.2.3. Data management and statistical analysis

Rates of clinical malaria and PP were aggregated at village level and linked to mortality data by village, year of study and age group. The age groups were defined as: 0-28 days (neonates),

1-11 months (infants), 1-4 years (child), 5-14 years (older child), 15-59 (adults) and 60+ (elderly). Crude and age specific all-cause/malaria-specific mortality rates were calculated by dividing the deaths in each group with the total person-years observed (pyo) in that group.

A measure of socioeconomic status was constructed based on household asset ownership using a composite score, derived from multiple correspondence analysis (MCA) (Amek et al., 2015) and categorized into 3 levels as least poor for the well off, poor for the average and poorest for the lowest rank while LLIN coverage was calculated as the percentage of households in a village owning at least one net per two people in a given year. Distance to health facilities was calculated as the networked distance of each household from the nearest health facility, and classified into 3 categories as less than 1km, 1 to 2km and greater than 2km. the elevation of each household was downloaded from the remote sensing United States geological survey (USGS) Earth Resources Observation and Science (EROS) website (NASA JPL).These variables were also aggregated at village level and linked to the parasitaemia and mortality data. The analysis considered two approaches; in one approach, the data were aggregated on a yearly basis, hence 9 years of observation; the second approach was aggregating the data into two periods (2007-2010 and 2011-2015).

For each age group, we fit Bayesian negative binomial geostatistical models on the mortality outcome to assess the relationship between PP and all-cause/malaria-specific mortality. Variable selection based on bivariate negative binomial models was used to identify potential confounders. Variables with a p-value below 0.1 were included in the final geostatistical models. Spatial correlation was taken into account by village specific random effects modelled via a Gaussian process with a mean of zero and an exponential correlation matrix of the distance between villages in the study (Diggle et al., 1998). Bayesian models were fitted in OpenBugs version 3.1.2 (Imperial College and Medical Research Council London, UK) using Markov Chain Monte Carlo (MCMC) simulation for parameter estimation. Regression coefficients from the Bayesian geostatistical model were exponentiated to obtain prevalence rate ratios (PRR)

and summarized by their posterior median and 95% Bayesian Credible Intervals (BCI). Covariate effects were considered statistically important when the BCI of the corresponding regression coefficients on the log scale did not include zero. Due to the nature of Bayesian statistical inference, we replace the terminology of statistically significant by statistically important effect when reporting our results. In this paper, we present the results for the association between clinical malaria and all-cause mortality; clinical malaria and malaria-specific mortality; and lastly PP and both all-cause/malaria-specific mortality in that order. Model formulation details are provided in Additional file 2.

4.3. **Results**

4.3.1. Descriptive statistics

Between the year 2007 and 2015, over 441,000 individuals were enrolled/monitored in the HDSS contributing a total of 2,114,223 pyo and 26,283 deaths, for an average crude death rate of 12.4 (95% Confidence Interval; 12.3-12.6) deaths per 1,000 pyo as shown in Table 4.1.

All-cause mortality during the study period rose from 15.5 (14.9-16.1) deaths per 1,000 pyo in 2007 to 18.8 (18.2-19.3) in 2008 then dropped to a low of 9.4 (9.0-9.8) in the year 2015 with malaria-specific mortality following a similar trend; rising from 1.3 (1.2-1.5) deaths per 1000 pyo in 2007 to a high of 3.5 (3.3-3.7) in 2008 but eventually dropping to 0.9 (0.7-1.0) deaths per 1000 pyo in 2015 (Table 4.1). The average PP during the whole study period was 35.8% (35.2-36.5); ranging between 27.3% in 2008 to a high of 39.7% in 2010 but then dropped over the years to 29.8% in 2015.

Table 4.1: All-cause/malaria specific mortality, clinical malaria and malaria parasite prevalence by year

Year	Person years of observation	Sampled population	Malaria parasite prevalence	Clinical malaria prevalence	All-cause mortality rate per 1,000 pyo*	Malaria-specific mortality rate per 1,000 pyo*
2007	181537	1270	29.6% (27.1-32.2)	7.2% (5.8-8.7)	15.5 (14.9-16.1)	1.3 (1.2-1.5)
2008	230374	1039	27.3% (24.6-30.2)	6.0% (4.6-7.6)	18.8 (18.2-19.3)	3.5 (3.3-3.7)
2009	230373	2508	39.0% (37.1-40.9)	7.5% (6.5-8.6)	15.6 (15.1-16.2)	2.9 (2.6-3.1)
2010	233871	5243	39.7% (38.4-41.0)	7.9% (7.2-8.6)	12.4 (11.9-12.8)	2.1 (1.9-2.3)
2011	238524	2091	39.2% (37.1-41.3)	8.5% (7.3-9.7)	10.9 (10.5-11.3)	1.4 (1.2-1.5)
2012	246254	2719	34.1% (32.3-35.9)	7.8% (6.8-8.8)	10.2 (9.8-10.6)	1.4 (1.2-1.5)
2013	249757	2358	34.5% (32.6-36.5)	10.5% (9.3-11.8)	10.5 (10.1-10.9)	1.6 (1.4-1.7)
2014	252173	1934	35.6% (33.4-37.8)	8.4% (7.2-9.7)	10.2 (9.8-10.6)	1.0 (0.9-1.2)
2015	251360	1756	29.8% (27.7-32.0)	7.5% (6.2-9.1)	9.4 (9.0-9.8)	0.9 (0.7-1.0)
Overall	**2114223**	**20918**	**35.8% (35.2-36.5)**	**8.1% (7.7-8.4)**	**12.4 (12.3-12.6)**	**1.8 (1.7-1.9)**

*pyo=person years of observation

A further breakdown of the positive slides showed that 8.1% (7.7-10.4) of the respondents had clinical malaria (parasites and fever); on average one fifth of all the positives had clinical malaria (Figure 4.1).

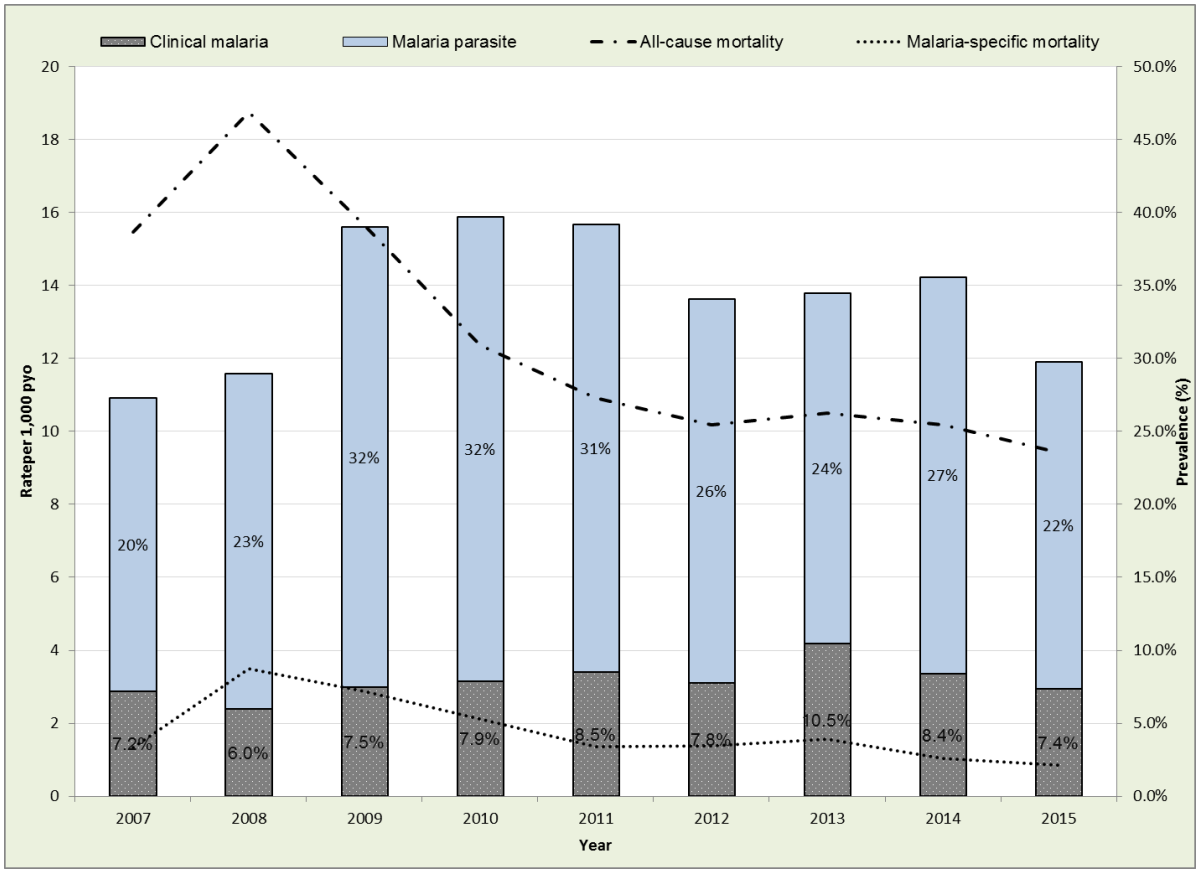
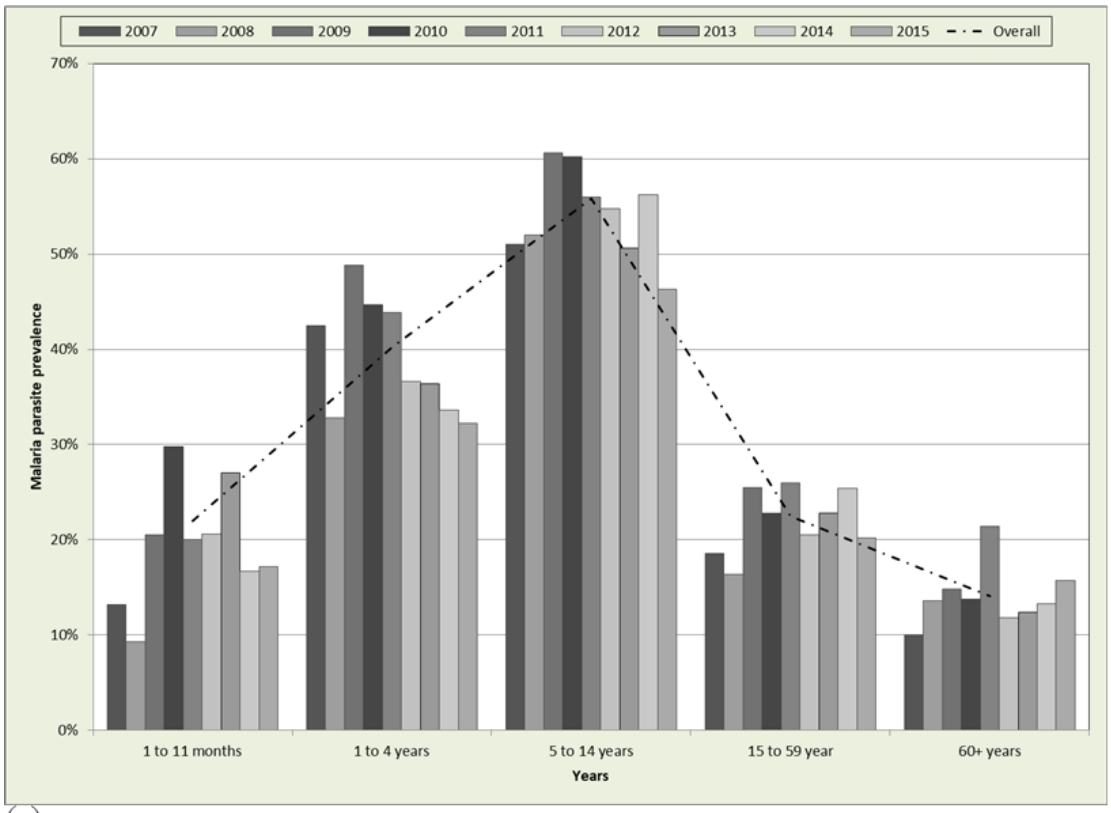
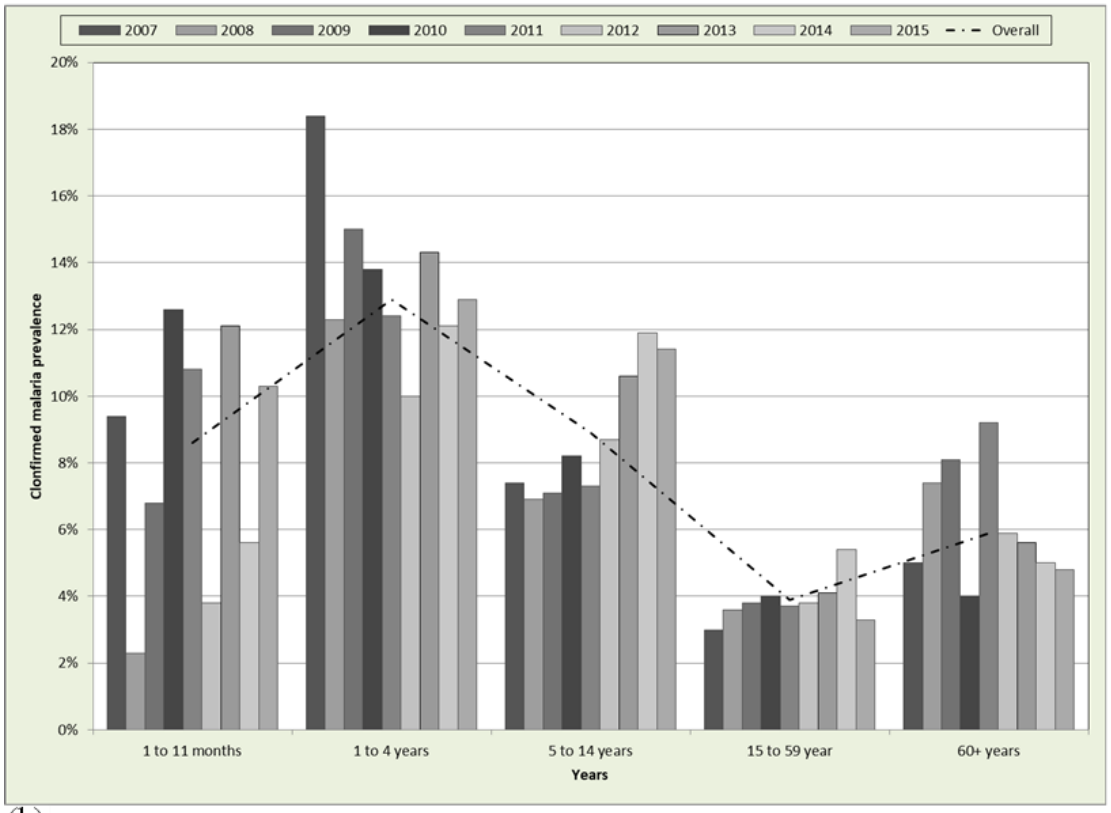


Figure 4.1: All-cause and malaria specific mortality rates versus malaria parasite and clinical malaria prevalence

The highest parasite prevalence was observed among older children aged 5-14 years, with an average PP of 56% (95% CI, 54-57), followed by children aged 1-4 years at 40% (39-41), adults at 22% (21-24), and infants at 22% (19-25); the elderly at 14% (12-16) had the lowest rate. The age distribution of prevalence indicates an increase in parasite prevalence from infanthood to older children followed by a drop as the population ages (Figure 4.2a). However, by including the presence of fever, we observed a rise in clinical malaria from infants to children aged 1 to 4 years after which it drops in the 5-14 age-group and in adults but rises slightly among the elderly. The highest prevalence of clinical malaria was in infants with a peak of 18.4% among those tested in the year 2007 (Figure 4.2b).



(a)



(b)

Figure 4.2: Malaria parasite prevalence (a) and clinical malaria (b) by age groups

4.3.2. Model based results

Relationship between clinical malaria and all-cause mortality

The following variables met the criteria for inclusion in the age-specific geostatistical mortality models: reported net usage, distance to health facilities, socioeconomic status, year of study and altitude. For comparability, these variables were included in the Bayesian models fitted by age group. Results in Table 4.2 show that the prevalence of confirmed malaria when aggregated over four and five-year periods, was associated with all-cause mortality in the combined age groups (RR=1.32; 95 % BCI: 1.01-1.74), in the 1–4-year-olds (RR=1.89; 1.00-3.51) and in the 5–14-year-olds (RR=3.94; 1.34-11.1). Increase in distance to health facilities was associated with higher mortality among neonates, children aged 1-4 years and the combined age group analysis. Risk of all-cause mortality was higher in the period 2007-2010 compared to 2011-2015 in all ages except in neonates. Higher SES and increased elevation were both associated with lower mortality. The association between reported net use and mortality was not statistically important across most age groups save for the elderly. The minimum distance at which spatial correlation was below 5% ranged from 13.2km to 50km for all the age groups. The analyses of the yearly prevalence data did not show a statistically important relation between confirmed malaria and all-cause mortality (see Additional file 4).

Table 4.2: Posterior estimates showing effects of prevalence (PP and confirmed malaria) on malaria-specific mortality†

Covariate	Neonates RR (95% BCI*)	Infants RR (95% BCI)	1-4 yrs. RR (95% BCI)	5-14 yrs. RR (95% BCI)	15-59 yrs. RR (95% BCI)	60 plus RR (95% BCI)	Overall** RR (95% BCI)
PP***	0.64 (0.29-1.34)	1.32 (0.86-2.04)	1.50 (0.98-2.23)	1.38 (0.69-2.69)	1.23 (0.94-1.57)	1.04 (0.78-1.35)	1.15 (0.96-1.37)
Clinical malaria***	1.05 (0.30-3.51)	1.65 (0.81-3.26)	1.89 (1.00-3.51)	3.94 (1.34-11.1)	0.99 (0.64-1.54)	1.07 (0.69-1.66)	1.32 (1.01-1.74)
Net use	1.18 (0.70-2.03)	1.06 (0.80-1.44)	0.82 (0.63-1.08)	0.97 (0.60-1.59)	0.88 (0.73-1.06)	0.78 (0.64-0.94)	0.91 (0.81-1.02)
Distance to health facility							
0 – 1 km	1	1	1	1	1	1	1
1 – 2 km	1.05 (0.81-1.37)	1.04 (0.90-1.20)	1.06 (0.93-1.22)	1.01 (0.80-1.28)	1.08 (0.99-1.18)	0.03 (0.94-1.13)	1.07 (1.01-1.14)
>2 km	1.33 (1.01-1.76)	1.06 (0.91-1.24)	1.16 (1.00-1.34)	1.28 (0.99-1.65)	1.09 (0.99-1.20)	1.07 (0.97-1.18)	1.12 (1.05-1.20)
SES							
Poorest	1	1	1	1	1	1	1
Poor	1.01 (0.78-1.29)	0.93 (0.80-1.07)	0.94 (0.82-1.07)	0.86 (0.68-1.08)	0.91 (0.83-0.99)	0.97 (0.88-1.06)	0.95 (0.90-1.01)
Least poor	0.83 (0.63-1.07)	0.85 (0.74-0.98)	0.84 (0.73-0.96)	0.91 (0.73-1.13)	0.88 (0.81-0.97)	0.95 (0.87-1.04)	0.89 (0.83-0.95)
Period							
2007-2010	1	1	1	1	1	1	1
2011-2015	1.05 (0.85-1.30)	0.57 (0.51-0.64)	0.60 (0.54-0.68)	0.77 (0.64-0.93)	0.68 (0.63-0.73)	0.96 (0.89-1.04)	0.71 (0.68-0.75)
Elevation							
1147 – 1243	1	1	1	1	1	1	1
1244 – 1293	1.13 (0.80-1.59)	1.09 (0.88-1.34)	1.15 (0.94-1.42)	0.81 (0.59-1.01)	1.03 (0.89-1.18)	0.98 (0.86-1.10)	1.01 (0.89-1.12)
1294 – 1327	0.84 (0.57-1.23)	0.80 (0.62-1.03)	0.97 (0.78-1.24)	0.98 (0.71-1.36)	0.99 (0.85-1.15)	0.92 (0.80-1.05)	0.87 (0.76-0.99)
1328 – 1365	1.17 (0.79-1.73)	0.82 (0.63-1.07)	0.99 (0.78-1.27)	0.96 (0.67-1.39)	0.93 (0.78-1.10)	0.96 (0.83-1.11)	0.89 (0.75-1.01)
>1365	0.92 (0.59-1.44)	0.61 (0.45-0.81)	0.80 (0.62-1.06)	1.09 (0.74-1.63)	0.92 (0.75-1.12)	0.98 (0.84-1.15)	0.80 (0.68-0.93)
Spatial Variance	0.49 (0.17-2.74)	0.86 (0.25-4.11)	0.77 (0.23-4.03)	0.51 (0.18-2.02)	0.57 (0.14-2.27)	0.34 (0.15-0.97)	0.96 (0.32-3.69)
Range$	24.07 (8.41-90.88)	15.25 (8.22-83.71)	17.38 (8.27-81.66)	23.59 (8.43-88.54)	19.77 (8.34-93.87)	50.09 (11.87-96.65)	13.23 (8.17-51.33)

†Mortality and malaria data aggregated by four to five-year periods (i.e. 2007-2010 and 2011-2015)
The effects are presented as the median of mortality rate ratios (RR) and 95% Bayesian credible intervals (BCI) adjusted for geographical variation and other predictors
** Age-adjusted
*** They are obtained from different models. Estimates of the rest of the predictors are from the models with confirmed malaria and do not differ from the PP model. PP estimates are only provided for comparison purposes
$ Minimum distance in kilometres at which spatial correlation is less than 5%

Relationship between clinical malaria and malaria-specific mortality

The pattern of association between clinical malaria and malaria-specific mortality across all age groups was similar to that of clinical malaria and all-cause mortality, however, the magnitude

of the estimates was higher. The effect of clinical malaria risk on malaria-specific mortality was statistically important and strong among children 5-14 years (RR=7.56; 1.20-39.54) and 1–4-year-olds (RR=3.37; 1.23-8.93). Meanwhile in the overall population, malaria-mortality rate increases two and half times for every increase in the proportion of clinical malaria by 1% (RR=2.50; 1.27-4.81) as shown in Table 4.3. Similar to all-cause mortality analysis, statistically important variables were elevation, distance to health facilities, year of study and socioeconomic status. Reported net use was not statistically important for malaria-specific mortality in any ages except among the elderly (RR=2.05; 1.04-4.34) in the yearly analysis, where we observed an elevated risk with higher levels of net use. The minimum distance at which spatial correlation was not important (<5%) ranged from 13.4 km to 50.42 km. The analyses of the yearly aggregated data did not show a statistically important relation between confirmed malaria risk and malaria-specific mortality (See Appendix 4)

Table 4.3: Posterior estimates for the effects of prevalence (PP and confirmed malaria) on malaria-specific mortality

Covariate	Infants RR (95% BCI)	1-4 yrs. RR (95% BCI)	5-14 yrs. RR (95% BCI)	15-59 yrs. RR (95% BCI)	60 plus RR (95% BCI)	Overall** RR (95% BCI)
PP***	1.73 (0.74-4.21)	2.29 (1.17-4.29)	0.56 (0.14-2.03)	1.55 (0.48-4.86)	2.24 (0.67-7.44)	1.66 (1.07-2.54)
Clinical malaria***	2.23 (0.55-8.36)	3.37 (1.23-8.93)	7.56 (1.20-39.54)	0.60 (0.08-3.91)	0.77 (0.09-5.64)	2.50 (1.27-4.81)
Net use	1.11 (0.61-1.97)	0.74 (0.48-1.14)	1.02 (0.40-2.35)	0.64 (0.28-1.35)	0.72 (0.30-1.77)	0.81 (0.61-1.11)
Distance to facility						
0 – 1 km	1	1	1	1	1	1
1 – 2 km	1.11 (0.85-1.48)	0.89 (0.72-1.11)	1.17 (0.75-1.86)	1.36 (0.93-2.03)	1.10 (0.71-1.72)	1.02 (0.89-1.17)
>2 km	1.19 (0.88-1.61)	1.00 (0.80-1.27)	1.56 (0.97-2.52)	1.13 (0.73-1.76)	1.15 (0.72-1.87)	1.09 (0.94-1.27)
SES						
Poorest	1	1	1	1	1	1
Poor	0.90 (0.67-1.19)	0.83 (0.67-1.03)	1.10 (0.72-1.65)	0.92 (0.61-1.35)	1.08 (0.70-1.63)	0.92 (0.80-1.06)
Least poor	0.96 (0.73-1.26)	0.78 (0.63-0.98)	0.94 (0.62-1.44)	0.86 (0.59-1.25)	1.05 (0.68-1.61)	0.88 (0.76-1.03)
Period						
2007-2010	1	1	1	1	1	1
2011-2015	0.58 (0.45-0.72)	0.56 (0.47-0.67)	0.72 (0.51-1.03)	0.64 (0.47-0.88)	0.68 (0.47-0.96)	0.56 (0.50-0.63)
Elevation						
1147 – 1243	1	1	1	1	1	1
1244 – 1293	1.05 (0.75-1.46)	1.17 (0.85-1.61)	0.94 (0.54-1.60)	1.05 (0.62-1.77)	0.89 (0.50-1.55)	1.08 (0.87-1.34)
1294 – 1327	0.47 (0.31-0.70)	1.10 (0.78-1.58)	1.02 (0.56-1.79)	0.79 (0.45-1.41)	1.00 (0.55-1.75)	0.84 (0.67-1.08)
1328 – 1365	0.65 (0.43-0.98)	1.00 (0.70-1.47)	0.99 (0.51-1.88)	0.89 (0.49-1.65)	1.15 (0.60-2.14)	0.88 (0.68-1.13)
>1365	0.36 (0.21-0.58)	0.88 (0.58-1.34)	1.39 (0.69-2.76)	0.58 (0.28-1.16)	1.29 (0.65-2.50)	0.71 (0.52-0.94)
Spatial variance	0.53 (0.17-2.43)	0.75 (0.21-4.40)	0.49 (0.16-2.43)	0.78 (0.18-4.08)	0.64 (0.22-2.65)	0.95 (0.21-7.62)
Range$	22.40 (8.42-92.52)	16.00 (8.22-74.12)	28.82 (8.47-91.99)	15.40 (8.21-86.65)	16.87 (8.32-75.36)	15.27 (8.19-78.32)

†Mortality and malaria data aggregated by four to five-year periods (i.e. 2007-2010 and 2011-2015)
The effects are presented as the median of mortality rate ratios (RR) and 95% Bayesian credible intervals (BCI) adjusted for geographical variation and other predictors
** Age-adjusted
*** They are obtained from different models. Estimates of the rest of the predictors are from the models with confirmed malaria and do not differ from the PP model. PP estimates are only provided for comparison purposes
$ Minimum distance in kilometres at which spatial correlation is less than 5%

Relationship between parasite prevalence and all-cause/malaria specific mortality

The relation between PP and all-cause mortality was not statistically important across all ages (Table 4.2). However, there was a statistically important association between PP and malaria-specific mortality (Table 4.3) among children aged 1-4 years (RR=2.29; 95% BCI: 1.17-4.29),

and in the combined age group (RR=1.66; 95% BCI: 1.07-2.54) when data was aggregated over 5 to 4 year. Analyses of yearly data did not reveal statistically important associations except between PP and all-cause mortality among the adults (RR=1.23; 95% BCI: 1.01-1.50) (See Additional file 4) and with malaria-specific mortality among the elderly (RR=3.42; 95 % CI: 1.39-8.63) (Appendix 5).

4.4. **Discussion**

Using data from community level cross-sectional surveys, our study shows that; parasite prevalence is associated with malaria-mortality in the overall population, while clinical malaria is associated with both all-cause and malaria-specific mortality more so in the age groups 1-4 years and 5-14years. This relationship was established by fitting over 50 different Bayesian geo-statistical models across different age groups on large data from verbal autopsies, longitudinal household surveys, and cross-sectional malaria parasitaemia surveys carried out annually over nine years in the HDSS located in a malaria endemic region of western Kenya. These data aggregated over four to five-year periods showed statistically important relations between clinical malaria and mortality (all-cause and malaria-specific) in the overall population, in children 1-4, and older children aged 5-14 years old, while PP had a statistically important association with malaria-specific mortality in 1–4-year-olds and in the overall population. Meanwhile, analyses of the same data, annually aggregated did not establish any association between prevalence of clinical malaria nor PP with either all-cause or malaria-specific mortality across most age groups except for all-cause mortality in adults aged 15-59 years and malaria-specific mortality in the elderly.

Studies in malaria-endemic areas have also shown that children above the age of 5 years are least affected by the malaria burden in terms of confirmed symptomatic malaria and mortality compared to other age groups, even though they remain the biggest reservoir of the malaria

parasites (Walldorf et al., 2015; Zhou et al., 2011). However, the long-term effects of declining transmission on mortality in this age group have not been well explored. This study showed a 7-fold increase in malaria-specific mortality for every 1% increase in clinical malaria prevalence, which was more than twice the effect in children 1–4-year-old. This finding could be attributed to low utilization of ITNs by older children compared to other age groups in this study as well as previous ones (Desai et al., 2014) or poor health care seeking behaviour reported in the same group (Bigogo et al., 2010), resulting in higher mortality rates when data is captured at household level compared to sentinel health facilities. This reinforces the importance of universal coverage of malaria control interventions particularly in high transmission areas.

The absence of an association between PP and all-cause mortality could be due to several factors. First, parasite prevalence from the community might capture more asymptomatic carriers who have acquired immunity from malaria disease, eventually recover without adverse outcomes and hence survive. Second, malaria mortality is usually preceded by severe illness, and therefore our PP data may be biased, as most of those who were severely ill may have gone to the hospital or succumbed to the disease prior to the time of the survey. Furthermore, an increase or decrease in mortality could be also due to other unmeasured factors that are unrelated to parasite prevalence; an example was shown by the influence of political instability on mortality in the year 2008 in Kisumu (Feikin et al., 2010) that resulted in massive disruption of health delivery.

The lack of association between PP or clinical malaria and mortality in the 15-49 age groups may be an indicator of misclassification of malaria as a cause of death by verbal autopsy. This weakness of verbal autopsy in identifying malaria as a cause of death among adults (Murray et al., 2012) could result in fewer deaths being classified as malaria than there really are in the population. Evidence suggests that people with HIV have more frequent episodes of symptomatic malaria (Whitworth et al., 2000) and that malaria increases HIV plasma viral load

and decreases CD4+ T cells (Alemu et al., 2013). Therefore, an alternative explanation could be that malaria specific mortality among adults may be classified by verbal autopsy as HIV/AIDS-related rather than malaria related.

The estimated effects of PP and clinical malaria were higher for malaria-specific mortality compared to all-cause. Furthermore, clinical malaria was a better predictor of mortality than PP. In fact, some of the asymptomatic infections may neither lead to severe disease nor death and therefore prevalence of clinical malaria is a better indicator for monitoring the disease burden at the population level. The stronger effect of clinical malaria and PP on malaria-specific mortality compared to all-cause mortality indicates that an increase in malaria transmission measures results in more malaria deaths which in turn inflate overall mortality. The stronger effect of prevalence on malaria specific mortality is because there is a clear biological cause and effect (Miller et al., 2002) and malaria infection can and does lead to mortality, however, the relationship between prevalence and all-cause mortality is diluted by other causes of mortality.

From our findings, it is worth noting that prevalence as a measure of transmission shows more stability in determining mortality over longer periods of time (4-5 years) compared to annual measures. Comparing estimates of the relation between mortality and malaria transmission measured by prevalence (of parasitaemia and confirmed malaria) in the current study, incidence (Khagayi et al., 2017) and EIR (Amek et al., 2018) in our previous studies, we note that incidence measured as slide positivity rate (SPR) followed by log EIR capture better the relationship between malaria transmission and mortality (Table 4.4). Confirmed malaria prevalence averaged over 4-5 years is likely to be more stable in areas of high transmission and therefore a useful measure of transmission over a long period while incidence and EIR capture the malaria-mortality relationship better over shorter periods (Amek et al., 2018; Khagayi et al., 2017). These differences could be due to the fact that PP one-off estimates can be misleading indicators of long-term transmission potential, since they vary markedly with

season (Drakeley et al., 2005) (Drakeley et al. 2005). These short-term fluctuation would then make it harder to associate yearly PP measures with mortality occurring all year round; suggesting that population based prevalence surveys do represent long term transmissions as opposed to short term changes.

Table 4.4: Comparison of estimates measuring the relation between malaria transmission and mortality from previous studies and current work

	Transmission measure	Neonates	Infants	Child	Older child	Adults	Elderly	All ages
(All-cause)	Log EIR[1]	**3.91 (3.53 - 4.32)***	**3.64 (3.40 - 3.89)**	**4.29 (3.89 - 4.73)**	-	-	-	-
	Slide Positivity Rate (Incidence)[2]	0.89 (0.13 - 5.70)	3.10 (0.36 - 13.12)	**4.29 (2.78 - 13.29)**	0.48 (0.15 - 2.05)	0.73 (0.39 - 1.42)	1.70 (0.79 - 4.45)	**1.55 (1.04 - 2.80)**
	Annual parasite prevalence[3]	1.23 (0.71 - 2.10)	1.04 (0.75 - 1.42)	1.11 (0.80 - 1.52)	1.15 (0.67 - 1.98)	**1.23 (1.01 - 1.50)**	1.04 (0.86 - 1.28)	1.10 (0.97 - 1.25)
	Annual clinical malaria prevalence[3]	1.03 (0.44 - 2.32)	1.22 (0.68 - 1.84)	1.38 (0.87 - 2.19)	1.97 (0.88 - 4.17)	1.16 (0.86 - 1.58)	0.99 (0.72 - 1.36)	1.16 (0.97 - 1.40)
	Five-year parasite prevalence[4]	0.64 (0.29 - 1.34)	1.32 (0.86 - 2.04)	1.50 (0.98 - 2.23)	1.38 (0.69 - 2.69)	1.23 (0.94 - 1.57)	1.04 (0.78 - 1.35)	1.15 (0.96 - 1.37)
	Five-year clinical malaria prevalence[4]	1.05 (0.30 - 3.51)	1.65 (0.81 - 3.26)	**1.89 (1.00 - 3.51)**	**3.94 (1.34 - 11.1)**	0.99 (0.64 - 1.54)	1.07 (0.69 - 1.66)	**1.32 (1.01 - 1.74)**
(Malaria-specific)	Log EIR[1]	-	**4.35 (3.72 - 4.95)**	**4.29 (3.61 - 5.06)**	-	-	-	-
	Slide Positivity Rate (Incidence)[2]	-	1.36 (0.23 - 9.85)	**9.48 (5.11 - 37.94)**	0.02 (0.003 - 0.33)	0.27 (0.02 - 3.24)	0.59 (0.01 - 13.15)	1.37 (0.51 - 3.73)
	Annual parasite prevalence[3]	-	1.31 (0.71 - 2.37)	1.50 (0.92 - 2.41)	0.80 (0.32 - 2.06)	0.78 (0.32 - 1.80)	**3.42 (1.39 - 8.63)**	1.31 (0.95 - 1.78)
	Annual clinical malaria prevalence[3]	-	0.94 (0.34 - 2.47)	1.58 (0.73 - 3.23)	3.65 (0.94 - 12.79)	0.36 (0.08 - 1.49)	1.67 (0.38 - 6.52)	1.34 (0.82 - 2.14)
	Five-year parasite prevalence[4]	-	1.73 (0.74 - 4.21)	**2.29 (1.17 - 4.29)**	0.56 (0.14 - 2.03)	1.55 (0.48 - 4.86)	2.24 (0.67 - 7.44)	**1.66 (1.07 - 2.54)**
	Five-year clinical malaria prevalence[4]	-	2.23 (0.55 - 8.36)	**3.37 (1.23 - 8.93)**	**7.56 (1.20 - 39.5)**	0.60 (0.08 - 3.91)	0.77 (0.09 - 5.64)	**2.50 (1.27 - 4.81)**

Estimates are Bayesian posterior medians and 95% Bayesian Credible Intervals (BCI)

Source

[1] Amek et al (2018). Infant and child mortality in relation to malaria transmission in KEMRI/CDC HDSS, Western Kenya: validation of verbal autopsy. Malar J. 17(1):37

[2] Khagayi et al (2017). Bayesian spatio-temporal modelling of mortality in relation to malaria incidence in Western Kenya. PLOS ONE.;12(7):e0180516

[3] Annually aggregated data. Current work.

[4] Data aggregated over 4-5 years. Current work.

*Statistical important effects are indicated in bold.

Higher socioeconomic status, shorter distance to health facilities and increasing altitude are known protective factors that were statistically important for both, all-cause and malaria-specific mortality. Individuals at a higher social status are more likely to live in well-constructed houses that offer better protection against endophagic/endophilic malaria vectors that transmit malaria in sub-Saharan Africa, afford better nutrition and pay for superior treatment (Sachs & Malaney, 2002). Increasing elevation is associated with lower temperatures which increase the development time of both vector and parasite (Githeko et al., 2006), resulting in lower transmission. Similarly, it has been shown that distance to health facilities influences mortality (Karra et al., 2016).

Lack of association between net use and mortality across ages except for yearly data among the elderly could be due to data aggregation at village level which diminished the expected individual level protection associated with net use reported in earlier studies during the 90's and early 2000's in the same region (Amek et al., 2018; Hawley et al., 2003). This change from earlier years could have been due to a number of factors among them ITN's having achieved maximum benefits, compromised effectiveness due to misuse/pyrethroid resistance or other unmeasured factors which countered their protective effect (Bayoh et al., 2014; Zhou et al., 2014). The diminished effect of net use might also be due to use of self-reported net use information which could lead to bias as it does not measure constant use. The negative effect of net use on malaria-specific mortality among the elderly, a group that has not been well researched in malaria cannot be explained adequately, and requires further investigation. We however hypothesize that since mortality is generally high in this age-group, at the same time society considers them vulnerable, issuance and use of ITNs could be higher and hence their protective effect is masked.

There are inherent limitations in survey data and in estimating malaria mortality using verbal autopsy that could influence our results. First, the surveys were conducted in specific months (i.e. in April, just before the rains (or just as they were starting) or June/July after the rains were

ending.); therefore, the prevalence estimates of may be biased by unexpected changes in climatic and environmental factors in other. Use of verbal autopsy as a tool for determining cause of death has been criticized (Murray et al., 2012), even though recent improvements in the InterVA coding have been said to reduce classification errors, especially at population level (Byass et al., 2015). Despite these limitations, the 9-year data in the study have been collected consistently in the same area using rigorous data collection methods and strict quality control measures. These data are thus unique in studying the relation between malaria prevalence and mortality across all groups in this population within a high endemic area.

4.5. Conclusion

Data from cross-sectional malaria prevalence were used to assess the relationship between malaria transmission and all-cause/malaria-specific mortality across age groups. The main findings were; i) prevalence as a measure of transmission is more stable over longer periods of time and its impact on mortality can be well investigated over periods of four to five years compared to incidence and EIR which better capture the malaria-mortality relationship on a yearly basis; ii) the risk of clinical malaria was strongly associated with mortality among 5-14 year olds; supporting the extension of control and prevention strategies to older children and adults; iii) symptomatic malaria from prevalence surveys at the population level can be used as a marker for increased malaria mortality. Therefore, strengthening health systems to capture high quality data on incidence would be useful and have greater relevance in predicting mortality. It is likely that the relationships between PP/clinical malaria and all-cause or malaria specific mortality are influenced by several factors such as the baseline endemicity of malaria, access to effective anti-malaria treatment and other factors. Therefore, analyses of data from other HDSS sites or similar platforms with differing levels of malaria endemicity different socio-economic status, or different access to effective anti-malarial drugs would be useful in understanding the contribution of PP to mortality across age groups.

Declarations

Ethical consideration

The study protocols for the KHDSS and parasite prevalence surveys were approved by the KEMRI Scientific and Ethics Review Unit (SERU) and the CDC Institutional Review Board (IRB). Informed written consent was obtained from the heads of compounds in conducting the household interviews while malaria prevalence data was collected from individuals who gave consent or from their parents/guardians if the participants were children.

Consent for publication

Not applicable.

Data availability

Data were obtained with permission of the Kisumu HDSS and Malaria branch steering committee. Any data requests may be sent to the respective steering committees, through Dr. Simon Kariuki (Skariuki@kemricdc.org) or Dr. Stephen Munga (Smunga@kemri.org).

Competing interests

The authors declare that there are no competing interests.

Funding

SK was supported by the European Research Council (ERC) IMCCA grant number 323180 and the Swiss National Foundation (SNF) program for Research on Global Issues for Development (R4D) project number IZ01Z0-147286 as part of his PhD work. The funders did not have any role in the study or its outcome.

Authors' contributions

SK MD and PV conceived the study; SK analysed the data and drafted the initial manuscript; PV supported data analysis; SK, MD, NA, VW, EO, CO, KO, LS, JG, FO, KL, AMS, SK, GB, SM, and MH designed the initial surveillance system, participated in data collection, processing and management. All authors gave inputs in revision of the first drafts, read and approved the final manuscript. The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the U.S. Centers for Disease Control and Prevention

Acknowledgement

We thank the HDSS community in Siaya County for the many years of support and participation in surveillance activities; the HDSS and Malaria branch staff of the KEMRI-CGHR for assisting in managing the platforms, data collection and management.

4.6. Appendix

Appendix 4.5.1: Study designs for the malaria survey data during 2007-2015

Month/Year of Study	Area(s)	Design and Context	Sampling design	Sample Size
April 2007	Whole HDSS area	HDSS	Systematic Random Sampling	1270
April 2008	Whole HDSS area	HDSS	Systematic random sampling	1039
April 2009	Whole HDSS area	HDSS	Cluster randomization (cluster unit = village) Villages selected by random sampling proportional to size	2508
April 2010	Whole HDSS area	HDSS	Systematic random sampling	5243
June-July 2011	Whole HDSS area	HDSS	Systematic random sampling	2091
June-July 2012	Whole HDSS area	HDSS	Systematic random sampling	2719
June-July 2013	Whole HDSS area	HDSS	Systematic random sampling	2358
2014	Whole HDSS area	HDSS	Systematic random sampling	1934
2015	Whole HDSS area	HDSS	Systematic random sampling	1756

Appendix 4.5.2: Bayesian model formulation

Let Y_{jt} be the observed number of deaths (all-cause or malaria-specific) in village *j* at time *t, j=1,2,...n* and *t=1,2,...9* for the model fitted on the annual data. We assume that Y_{jt} arises from a negative binomial distribution, $Y_{jt} \sim NB(\mu_{jt}, r)$ where μ_{jt}, is the expected number of deaths and *r* is the dispersion parameter. We modelled malaria prevalence and other covariates (***X***) by on the log scale of μ_{jt}

$$log(\mu_{jt}) = log(N_{jt}) + \beta_0 + \sum_1^k \beta_k X_{jtk} + \phi_j \qquad , k=1,2,...K$$

where N_{jt} the total person time contributed by persons in village *j* at time *t* (in person years, py), $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_K)^T$ are the regression coefficients (malaria prevalence and other predictors) and ϕ_j the village specific spatial effects. We assumed that $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots \phi_n)^T$ are modeled by a Gaussian process, that is $\boldsymbol{\phi} \sim MVN(0, \sigma_1^2 R)$ and that R is an exponential correlation matrix of the distance between villages, i.e. $R_{ij} = \exp(-\rho d_{kl})$ where d_{kl} is the distance between villages k and l, ρ is the rate of *correlation* decay with distance. The minimum distance at which the spatial correlation is less than 5% is called effective range and is defined by the value of $3/\rho$ (1). The σ_1^2 is the variance of the spatial process. We specified non-informative normal prior distributions with mean zero and large variance for the regression coefficients $\beta_i \sim N(0, 10^3)$, an inverse gamma prior for σ_1^2, that is $\sigma_1^2 \sim IG(2.01, 1.01)$ and a Uniform prior distribution for ρ, that is $\rho \sim U(a, b)$, where a and b are chosen such as the effective range is within the maximum and minimum distances of the village's locations.

The Bayesian models were then fitted using Markov Chain Monte Carlo (MCMC) simulation algorithm in OpenBugs version 3.1.2 (Imperial College and Medical Council, London, UK) to estimate model parameters (Gelfand & Smith, 1990). We ran a two-chain sampler for 100,000 iterations, discarding the first 10,000 iterations. Convergence was assessed by the Gelman-Rubin diagnostic (Gelman & Rubin, 1992).

Appendix 4.5.3: Posterior estimates of the effects of prevalence on all-cause mortality aggregated annually. †

Covariate	Neonates RR (95% CI*)	Infants RR (95% CI)	1-4 yrs. RR (95% CI)	5-14 yrs. RR (95% CI)	15-59 yrs. RR (95% CI)	60 plus RR (95% CI)	Overall** RR (95% CI)
PP***	1.23 (0.71-2.10)	1.04 (0.75-1.42)	1.11 (0.80-1.52)	1.15 (0.67-1.98)	1.23 (1.01-1.50)	1.04 (0.86-1.28)	1.1 (0.97-1.25)
Clinical malaria***	1.03 (0.44-2.32)	1.22 (0.68-1.84)	1.38 (0.87-2.19)	1.97 (0.88-4.17)	1.16 (0.86-1.58)	0.99 (0.72-1.36)	1.16 (0.97-1.40)
Net use	1.3 (0.87-1.93)	0.85 (0.68-1.07)	0.87 (0.70-1.07)	0.77 (0.52-1.11)	1.07 (0.93-1.22)	0.9 (0.77-1.04)	0.96 (0.88-1.04)
Distance to facility							
0 – 1 km	1	1	1	1	1	1	1
1 – 2 km	1.06 (0.82-1.95)	1.08 (0.93-1.24)	1.09 (0.95-1.26)	1.05 (0.83-1.33)	1.1 (1.01-1.22)	1.04 (0.95-1.14)	1.09 (1.03-1.16)
>2 km	1.4 (1.06-2.71)	1.1 (0.94-1.29)	1.18 (1.01-1.37)	1.35 (1.05-1.74)	1.1 (1.00-1.22)	1.08 (0.97-1.19)	1.15 (1.07-1.22)
SES							
Poorest	1	1	1	1	1	1	1
Poor	1.09 (0.52-1.20)	0.95 (0.82-1.09)	0.94 (0.82-1.08)	1.06 (0.84-1.35)	0.97 (0.88-1.06)	0.93 (0.85-1.02)	0.98 (0.92-1.03)
Least poor	0.93 (0.35-0.84)	0.92 (0.79-1.07)	0.81 (0.70-0.94)	1.02 (0.80-1.32)	0.94 (0.86-1.03)	0.95 (0.86-1.05)	0.92 (0.86-0.99)
Year							
2007	1	1	1	1	1	1	1
2008	1.67 (1.14-2.45)	1.57 (1.28-1.94)	1.59 (1.28-2.00)	1.6 (1.03-2.54)	1.07 (0.93-1.23)	1.14 (0.96-1.35)	1.3 (1.19-1.42)
2009	1.29 (0.82-2.02)	1.25 (0.97-1.61)	1.45 (1.13-1.88)	1.43 (0.87-2.36)	0.93 (0.80-1.09)	1.1 (0.91-1.33)	1.11 (1.00-1.23)
2010	0.58 (0.39-0.86)	0.9 (0.74-1.11)	0.97 (0.80-1.20)	1.27 (0.87-1.95)	0.71 (0.63-0.80)	1 (0.87-1.16)	0.84 (0.78-0.91)
2011	0.64 (0.41-0.98)	0.77 (0.61-0.97)	0.8 (0.63-1.01)	0.91 (0.58-1.46)	0.63 (0.55-0.72)	1.01 (0.86-1.18)	0.75 (0.68-0.82)
2012	1.14 (0.79-1.64)	0.57 (0.45-0.72)	0.66 (0.49-0.87)	1.21 (0.81-1.89)	0.59 (0.52-.067)	1.01 (0.87-1.18)	0.71 (0.65-0.77)
2013	1.1 (0.77-1.58)	0.73 (0.59-0.91)	0.77 (0.62-0.96)	1.02 (0.68-1.58)	0.55 (0.49-0.63)	0.96 (0.83-1.12)	0.71 (0.65-0.77)
2014	1.24 (0.80-1.91)	0.66 (0.49-0.88)	0.66 (0.49-0.87)	0.99 (0.59-1.66)	0.6 (0.51-0.70)	0.97 (0.81-1.16)	0.69 (0.62-0.77)
2015	0.86 (0.50-1.42)	0.52 (0.38-0.72)	0.52 (0.38-0.69)	1.28 (0.82-2.08)	0.5 (0.42-0.59)	1.02 (0.86-1.21)	0.63 (0.57-0.70)
Elevation							
1147 – 1243	1	1	1	1	1	1	1
1244 – 1293	1.16 (0.85-1.61)	1.07 (0.86-1.33)	1.15 (0.95-1.42)	0.83 (0.60-1.15)	1.05 (0.90-1.22)	0.99 (0.86-1.12)	1.01 (0.90-1.11)
1294 – 1327	0.87 (0.61-1.23)	0.77 (0.59-0.98)	1.01 (0.81-1.26)	0.96 (0.68-1.35)	1.03 (0.87-1.22)	0.93 (0.79-1.07)	0.89 (0.79-1.00)
1328 – 1365	1.23 (0.85-1.77)	0.79 (0.61-1.03)	1.02 (0.80-1.30)	0.99 (0.67-1.44)	0.96 (0.79-1.15)	0.97 (0.83-1.14)	0.9 (0.78-1.02)
>1365	0.99 (0.65-1.52)	0.54 (0.40-0.73)	0.81 (0.61-1.07)	1.1 (0.71-1.67)	0.96 (0.78-1.20)	1 (0.83-1.19)	0.82 (0.70-0.95)
Spatial Variance	0.06 (0.03-0.33)	0.23 (0.05-0.37)	0.19 (0.04-0)	0.17 (0.04-0.36)	0.19 (0.05-0.36)	0.08 (0.03-0.33)	0.24 (0.06-0.37)
Range[$]	54.34 (9.05-98.08)	13.2 (8.17-55.31)	15.53 (8.20-84.11)	17.69 (8.29-78.62)	15.54 (8.24-65.62)	38.26 (9.14-95.12)	12.39 (8.13-50.99)

The effects are presented as the median of mortality rate ratios (RR) and 95% Bayesian credible intervals (BCI) adjusted for geographical variation and other predictors

** Age-adjusted

*** They are obtained from different models. Estimates of the rest of the predictors are from the models with confirmed malaria and do not differ from the PP model. PP estimates are only provided for comparison purposes

$ Minimum distance in kilometres at which spatial correlation is less than 5%

Appendix 4.5.4: All-cause mortality, malaria-specific mortality and clinical malaria/parasite prevalence by year and age group

Year	Indicators	All	Neonates	Infants	Child	Older child	Adult	Elderly
2007	**Person years of observation**	181537	521	6222	22379	49372	83225	14059
	All-cause death rate	15.5	243.7	59.3	17.1	2.0	13.4	50.6
	Malaria death rate	1.3	-	9.0	3.8	0.3	0.6	1.1
	Sampled population	1270	-	53	228	363	526	100
	PP (clinical malaria)	30% (7%)	-	9% (9%)	43% (18%)	51% (7%)	19% (3%)	10% (5%)
2008	**Person years of observation**	230374	679	8203	29455	63830	107040	17498
	All-cause death rate	18.8	328.5	86.6	28.1	2.4	13.5	54.9
	Malaria death rate	3.5	-	26.3	12.5	0.5	0.7	2.4
	Sampled population	1039	-	34	171	304	440	81
	PP (clinical malaria)	27% (6%)	-	9% (2%)	33% (12%)	52% (7%)	16% (4%)	14% (7%)
2009	**Person years of observation**	230373	609	7666	29535	64664	106636	17915
	All-cause death rate	15.6	246.3	68.2	22.0	2.3	11.3	51.5
	Malaria death rate	2.9	-	20.2	8.8	0.8	0.8	2.9
	Sampled population	2508	-	117	521	708	939	223
	PP (clinical malaria)	39% (8%)	-	21% (7%)	49% (15%)	61% (7%)	25% (4%)	15% (8%)
2010	**Person years of observation**	233871	575	7165	29773	66475	108038	18316
	All-cause death rate	12.4	132.2	51.1	16.2	2.0	9.1	46.5
	Malaria death rate	2.1	-	15.1	6.4	0.8	0.4	2.7
	Sampled population	5243	-	191	1177	1717	1861	297
	PP (clinical malaria)	40% (8%)	-	30% (13%)	45% (14%)	60% (8%)	23% (4%)	14% (4%)
2011	**Person years of observation**	238524	550	6817	29790	68937	111025	18393
	All-cause death rate	10.9	130.9	44.3	13.2	1.6	7.8	47.0
	Malaria death rate	1.4	-	11.0	4.7	0.4	0.3	1.3
	Sampled population	2091	-	130	776	495	592	98
	PP (clinical malaria)	39% (8%)	-	20% (11%)	44% (12%)	56% (7%)	26% (4%)	21% (9%)
2012	**Person years of observation**	246254	575	6863	30021	72024	115190	18871
	All-cause death rate	10.2	219.2	32.5	11.0	1.8	7.2	45.8
	Malaria death rate	1.4	-	9.18	4.13	0.68	0.36	1.38
	Sampled population	2719	-	209	1336	473	599	102
	PP (clinical malaria)	34% (8%)	-	21% (4%)	37% (10%)	55% (9%)	21% (4%)	12% (6%)
2013	**Person years of observation**	249757	584	7146	29858	74793	118096	19669
	All-cause death rate	10.5	256.8	40.9	12.6	1.7	6.8	45.4
	Malaria death rate	1.6	-	11.8	4.6	0.6	0.4	2.6
	Sampled population	2358	-	141	1044	492	592	89
	PP (clinical malaria)	34% (10%)	-	27% (12%)	36% (14%)	51% (11%)	23% (4%)	12% (6%)
2014	**Person years of observation**	252173	519	6778	29289	75891	119903	20305
	All-cause death rate	10.2	275.5	34.7	11.7	1.6	7.2	44.2
	Malaria death rate	1.0	-	7.5	3.3	0.3	0.4	1.5
	Sampled population	1934	-	54	232	664	804	180
	PP (clinical malaria)	36% (8%)	-	17% (6%)	34% (12%)	56% (12%)	25% (5%)	13% (5%)
2015	**Person years of observation**	251360	448	5979	27917	75876	120924	20695
	All-cause death rate	9.4	185.4	26.6	9.1	1.9	6.6	46.2
	Malaria death rate	0.9	-	3.2	1.9	0.3	0.1	1.0
	Sampled population	1756	-	29	227	570	783	147
	PP (clinical malaria)	30% (7%)	-	17% (10%)	32% (13%)	46% (11%)	20% (3%)	16% (5%)

Appendix 4.5.5: Posterior estimates of the effects of prevalence on malaria-specific mortality aggregated annually

Covariate		Infants RR (95% BCI*)	1-4 yrs. RR (95% BCI)	5-14 yrs. RR (95% BCI)	15-59 yrs. RR (95% BCI)	60 plus RR (95% BCI)	Overall** RR (95% BCI)
PP***		1.31 (0.71-2.37)	1.50 (0.92-2.41)	0.8 (0.32-2.06)	0.78 (0.32-1.80)	3.42 (1.39-8.63)	1.31 (0.95-1.78)
Clinical malaria***		0.94 (0.34-2.47)	1.58 (0.73-3.23)	3.65 (0.94-12.79)	0.36 (0.08-1.49)	1.67 (0.38-6.52)	1.34 (0.82-2.14)
Net use		1.08 (0.68-1.65)	0.94 (0.67-1.31)	0.64 (0.33-1.27)	0.75 (0.40-1.42)	2.05 (1.04-4.34)	0.97 (0.79-1.21)
Distance to facility							
	0 – 1 km	1	1	1	1	1	1
	1 – 2 km	1.13 (0.86-1.50)	0.93 (0.76-1.15)	1.23 (0.78-1.98)	1.38 (0.94-2.07)	1.04 (0.67-1.63)	1.06 (0.92-1.22)
	>2 km	1.16 (0.86-1.57)	1.03 (0.82-1.30)	1.61 (1.00-2.65)	1.14 (0.74-1.79)	1.11 (0.68-1.81)	1.11 (0.94-1.22)
SES							
	Poorest	1	1	1	1	1	1
	Poor	0.89 (0.67-1.19)	0.94 (0.75-1.16)	0.84 (0.55-1.28)	0.81 (0.55-1.19)	0.98 (0.63-1.53)	0.89 (0.78-1.03)
	Least poor	0.87 (0.64-1.18)	0.81 (0.64-1.03)	0.71 (0.45-1.13)	0.77 (0.51-1.15)	1.18 (0.74-1.88)	0.84 (0.72-0.98)
Year	2007	1	1	1	1	1	1
	2008	3.05 (1.94-4.94)	3.37 (2.32-5.02)	2.09 (0.71-6.82)	1.36 (0.73-2.57)	2.44 (1.01-6.23)	2.74 (2.14-3.48)
	2009	1.93 (1.11-3.38)	2.37 (1.54-3.69)	4.76 (1.74-15.02)	2.06 (1.07-3.96)	1.35 (0.47-3.84)	2.17 (1.66-2.86)
	2010	1.72 (1.11-2.76)	1.69 (1.18-2.48)	3.72 (1.56-10.63)	0.9 (0.52-1.62)	2.52 (1.17-5.82)	1.66 (1.32-2.11)
	2011	1.3 (0.78-2.24)	1.35 (0.91-2.06)	1.82 (0.65-5.75)	0.73 (0.37-1.43)	1.12 (0.45-2.84)	1.14 (0.88-1.48)
	2012	1.15 (0.70-1.94)	1.08 (0.72-1.63)	3.07 (1.22-9.07)	0.67 (0.36-1.27)	1.13 (0.47-2.87)	1.05 (0.81-1.35)
	2013	1.34 (0.83-2.22)	1.3 (0.88-1.94)	2.72 (1.09-7.86)	0.97 (0.56-1.77)	1.81 (0.82-4.33)	1.25 (0.99-1.60)
	2014	0.78 (0.39-1.52)	0.82 (0.48-1.38)	1.57 (0.47-5.40)	0.79 (0.37-1.63)	1.08 (0.39-3.05)	0.79 (0.57-1.08)
	2015	0.19 (0.01-0.56)	0.51 (0.27-0.93)	1.24 (0.35-4.37)	0.16 (0.03-0.50)	0.41 (0.05-1.53)	0.36 (0.23-0.54)
Elevation							
	1147 – 1243	1	1	1	1	1	1
	1244 – 1293	1.07 (0.76-1.50)	1.19 (0.88-1.61)	0.89 (0.51-1.53)	1.05 (0.62-1.77)	0.9 (0.49-1.63)	1.1 (0.88-1.37)
	1294 – 1327	0.49 (0.32-0.75)	1.23 (0.88-1.68)	0.97 (0.55-1.72)	0.81 (0.47-1.41)	1.05 (0.55-1.98)	0.9 (0.69-1.17)
	1328 – 1365	0.72 (0.47-1.11)	1.13 (0.78-1.58)	0.98 (0.51-1.86)	0.96 (0.54-1.71)	1.12 (0.56-2.25)	0.95 (0.72-1.23)
	>1365	0.4 (0.23-0.66)	0.94 (0.62-1.38)	1.28 (0.64-2.54)	0.61 (0.31-1.19)	1.21 (0.56-2.61)	0.75 (0.55-1.00)
Spatial Variance		0.19 (0.04-0.36)	0.16 (0.04-0.36)	0.06 (0.03-0.29)	0.14 (0.03-0.36)	0.22 (0.05-0.37)	0.22 (0.05-0.37)
Range[$]		15.91 (8.23-74.54)	18.22 (8.25-84.10)	50.42 (10.35-97.07)	20.75 (8.39-90.56)	13.44 (8.18-60.19)	13.36 (8.17-63.63)

The effects are presented as the median of mortality rate ratios (RR) and 95% Bayesian credible intervals (BCI) adjusted for geographical variation and other predictors

** Age-adjusted

*** They are obtained from different models. Estimates of the rest of the predictors are from the models with confirmed malaria and do not differ from the PP model. PP estimates are only provided for comparison purposes

$ Minimum distance in kilometres at which spatial correlation is less than 5%

5. Chapter 5: Bayesian geo-statistical modelling of population level anaemia burden in relation to malaria and parasitic worms' prevalence across age groups in western Kenya.

Sammy Khagayi[1,2,3], Christos Kokliaris[2,3], Brian Nyawanda[1,2,3], Meghna Desai[4], Daniel Kwaro[1], Stephen Munga[1], Kephas Otieno[1], Mary Hamel[4], Laurence Slutsker[4], John Gimnig[4], Aaron Samuels[4,5], Jürg Utzinger[2,3], Simon Kariuki[1], and Penelope Vounatsou[2,3]*

1. Kenya Medical Research Institute-Centre for Global Health Research, Kisumu, Kenya
2. Swiss Tropical and Public Health Institute, Basel, Switzerland
3. University of Basel, Basel, Switzerland
4. Centers for Disease Control and Prevention, Center for Global Health, Division of Parasitic Diseases and Malaria, Malaria branch, Atlanta, GA, USA
5. Centers for Disease Control and Prevention, Kisumu, Kenya

[§]Corresponding author

This manuscript is prepared for submission to Plos Global Health

Abstract

Background

Anaemia remains a leading cause of morbidity and mortality worldwide with a disproportionate burden in low and middle-income countries. However, previous studies of anaemia epidemiology have been geographically limited with little detail about severity or aetiology across different age groups in the same population.

Methods

We collated malaria parasitaemia data from annual cross-sectional surveys between 2007 and 2015 together with Schistosomiasis/helminths data from geo-statistical models aggregated at village level. Using these, we fitted Bayesian geo-statistical regression models to investigate the association between malaria and parasitic worms' prevalence with anaemia risk across different age groups.

Results

We established a persistently high prevalence of anaemia affecting 42.3% (CI: 41.2-43.4%) of the population, which is classified as a severe public health indicator. Across all age groups, our analysis showed on average, a consistently high prevalence of anaemia across the years, mostly affecting young children 1-11 months at 67.1% (CI: 62.5-71.6%), 1-4 years at 61.0% (CI: 58.9-63.1%); the elderly at 47.1% (CI: 43.6-50.6%) and women at 45.3% (CI: 43.8-46.8%). Malaria parasitaemia, clinical malaria, age, sex and socioeconomic status as a proxy for malnutrition were significantly associated with a higher anaemia risk and lower levels of Haemoglobin (Hb) concentration. Schistosomiasis and helminths risk were not associated with anaemia prevalence in this study area.

Conclusions

The high, stable anaemia prevalence across all age groups and the multifaceted nature of factors determining anaemia prevalence underscore the need for an integrated approach to reduce its burden in this and other populations with endemic malaria and similar health challenges.

Keywords: Anaemia, Malaria, Soil transmitted helminths, Schistosomiasis, Bayesian spatio-temporal, Health and Demographic Surveillance system.

5.1. Introduction

Despite its decline in recent years, anaemia remains a major public health problem affecting over 1.9 billion people with the biggest burden experienced in Sub-Saharan Africa (SSA) (GBD Anaemia Collaborators, 2023; Kassebaum, 2016; Safiri et al., 2021; Stevens et al., 2013). In Kenya, available data estimates anaemia prevalence of 40-50% and 20-40% in children less than 5 years and women of reproductive age respectively (Division of National Malaria Programme (DNMP) & ICF, 2021; World Health Organization, 2015a). In the

Anaemia disproportionately affects the poorest, marginalised groups and is associated with a wide range of poor health outcomes across all age groups (Balarajan et al., 2011). It increases the risk of maternal and child mortality/morbidity, promotes poor birth outcomes, increases chances of lifelong conditions, exacerbates existing or acquired health conditions and impairs normal bodily functions (Anand, 2008; Balarajan et al., 2011; Cavill et al., 2006; Haider et al., 2013; Terekeci et al., 2010). The aetiology of anaemia varies from place to place, with iron deficiency identified as the main cause in SSA. Its manifestation, is also complex in low-income countries due to the presence of other factors, among them infectious diseases such as malaria, parasitic worms, HIV/AIDS, malnutrition and genetic disorders (Ejigu et al., 2018; Foote et al., 2013; Kassebaum, 2016; Soares Magalhães & Clements, 2011).

Malaria and parasitic worms, especially Schistosoma *mansoni* and soil transmitted helminths STH (Hookworms, Ascaris *lumbricoides* and Trichuris *trichiura*) have been shown in different studies to be among the major contributors of anaemia (GBD Anaemia Collaborators, 2023; White, 2018). Malaria parasites are known to cause haemolysis and disrupt the normal development of red blood cells (Menendez et al., 2000). Soil transmitted helminths burrow their teeth into the digestive system mucosa, feed on host tissues and cause blood loss, leading to anaemia (Caldrer et al., 2022). While schistosomiasis causes anaemia through a combination of effects, including blood loss, red blood cell destruction in the spleen, immune mechanisms, hence iron deficiency, and general inflammation (Butler et al., 2012).

Due to the complex malaria-anaemia relationship, some studies recommend reporting them as one disease (Hamel et al., 2011; Ondeto et al., 2022; Sewe et al., 2015). In some instances, anaemia is used to estimate the burden of malaria (White, 2018). Furthermore, regions with high malaria prevalence also tend to have a high burden of schistosomiasis and soil-transmitted helminths (STH) (Foote et al., 2013; Odiere et al., 2011; Wiegand et al., 2017).

One significant aspect of malaria, helminths, schistosomiasis and other anaemia related environmental/climatic factors is that they are spatially correlated. Which means that nearby neighbourhoods could be very similar to each other compared to those at a distance in terms of infection spread (Ejigu et al., 2018). This spatial relationship is vital to consider when studying anaemia spread and its associations.

While global studies have estimated the burden of anaemia across all age groups, (GBD Anaemia Collaborators, 2023; Stevens et al., 2013; World Health Organization, 2015a). local studies have often focused on specific age brackets, such as children under 5 years (Desai et al., 2005; Foote et al., 2013), women of reproductive age (Ouma et al., 2007) or school going children (Koukounari et al., 2008). However, there is a lack of comprehensive research profiling of anaemia across different age groups.

Considering the various factors at play, it is crucial to pinpoint the exact impact of malaria on anaemia prevalence. In this study, we profile the burden of anaemia, and explores its association with malaria, STH, and schistosomiasis using Bayesian geo-statistical models in a health and demographic surveillance system (HDSS) located in western Kenya.

5.2. Materials and methods

5.2.1. Study area and population

This study was carried out in the KEMRI health and demographic surveillance system (KHDSS), a longitudinal follow up of all residents in a specified geographical area. Details of

the KHDSS are published elsewhere. Briefly, the KHDSS follows a population of over 280,000 people in 393 villages located in Siaya in western Kenya covering 700km^2. (Odhiambo et al., 2012). It is located in a region with high malaria prevalence, year round transmission (Division of National Malaria Programme (DNMP) & ICF, 2021; Hamel et al., 2011; Ondeto et al., 2022), and a relatively high prevalence of both schistosomiasis and STH (Odiere et al., 2011; Wiegand et al., 2017).

5.2.2. Haemoglobin levels and malaria data

Annual parasitaemia surveys were carried out between 2007 and 2015 at randomly selected compounds within the HDSS to obtain individual level haemoglobin concentration and malaria parasite presence. The in-depth sampling and sample size determination have been described in elsewhere (Were et al., 2018). In brief, households were selected by systematic random sampling and stratified by sub-regions of the HDSS. Starting from a randomly chosen number in an ordered list of compounds for each region, households were systematically chosen until the desired sample size was met.

Trained interviewers then visited these compounds, and administered a structured questionnaire to collect information on demographics, risk factors for malaria infection, healthcare seeking, previous illness, socioeconomic status and history of fever. The participant's finger was pricked to obtain a blood specimen from all individuals present in the sampled households. Haemoglobin concentration levels were measured using a portable photometer (HemoCue®, Ängelholm, Sweden). Thick and thin blood smear slides were prepared for malaria microscopy, stained with a 10% Giemsa and examined for parasite presence by expert microscopists.

Using the World Health Organizations (WHO) guidelines for anaemia classification (World Health Organization, 2011), we classified the respondents as anaemic or not based on the cut-offs below (Table 5.1). Since we did not have the pregnancy status of women, we classified all

women above 15 years in the same category. Likewise, because all study locations were on average at elevations below 1,000m and did not vary widely, we did not apply altitude correction as recommended

Table 5.1: WHO Haemoglobin thresholds for defining anaemia

Age or gender	Hb threshold (g/dL)
Children (Amek, 2013)	11.0
Children (5-12 years)	11.5
Older children (12-15 years)	12.0
Women (>15 years)	12.0
Men (>15 years)	13.0

From the microscopy results, the participants were classified into either parasitaemia positive or negative. Clinical malaria was defined as having positive malaria parasitaemia together with either reported fever in the last 24 hours, or a temperature of 37.5^0 and above.

5.2.3. Helminths and Schistosomiasis data

Helminths data was derived from prevalence estimates of ascaris, trichurias, and hookworm infection extracted from high resolution geostatistical maps derived using Bayesian hierarchical models across Sub-Saharan Africa (Kokaliaris et al., 2022). In brief, hierarchical models were used to predict the prevalence of schistosomiasis and STH by relating cross-sectional schistosomiasis/STH survey data with socioeconomic and environmental predictors. We extracted the STH and schistosomiasis prevalence estimates to the household level from Kokaliaris's spatial maps using geocoded HDSS locations. Since these data was predicted for two periods (<2011 and ≥2011), we allocated the Helminths and Schistosomiasis estimates from pre-2011 period to the years between 2007 and 2010 and the post-2011 period to the years between 2011 and 2015.

5.2.4. Socioeconomic data

Socioeconomic status (SES) is a significant determinant of nutritional well-being, especially in areas with limited resources (Fotso, 2007; Müller & Krawinkel, 2005; World Health Organization, 2017a). People at the lower end of the SES spectrum often face challenges like inadequate diets, restricted access to formal healthcare, and exposure to unsanitary environments. This results in micronutrient deficiencies and a higher risk of infections. Consequently, SES was chosen as an indirect measure to assess the impact of malnutrition on anaemia. To determine SES, household assets were used to construct a composite score through the multiple correspondence analysis (MCA) method. This score was then divided into three categories: 'least poor' for the affluent, 'poor' for the middle group, and 'poorest' for those at the bottom, as outlined in previously (Amek et al., 2015).

5.2.5. Data management and statistical analysis

Data from annual parasitaemia surveys were combined into a single file, which was then cleaned and matched with socioeconomic status (SES) data from households. This linkage was achieved using unique identifiers for individuals and households. Using the households' GPS locations, we extracted STH and schistosomiasis prevalence estimates from geostatistical maps produced by Kokaliaris et al (Kokaliaris et al., 2022). Schistosomiasis and Helminthiasis prevalence were treated as continuous variables.

Age was calculated by subtracting the date of birth from the interview date and grouped into 5 categories: infants (1-11 months), children (1-4 years), older children (5-14 years), adults (15-59 years), and the elderly (60 years and above). Average haemoglobin levels and anaemia prevalence were assumed to follow a Gaussian distribution, similar to the overall population. In this paper we present the adjusted average haemoglobin (Hb) levels and the prevalence of anaemia for the population.

Initial data handling, bivariate analysis, and choice of variables for the final model were conducted using Stata version 14.1 (Stata Corporation USA). Only variables with a p-value less than 0.1 were incorporated into the final geostatistical models. After the initial variable selection, we included the following covariates in the final Bayesian geostatistical models; malaria parasite presence/clinical malaria, sex, year of study, Schistosomiasis prevalence, soil transmitted helminth prevalence, elevation, and socio-economic status to investigate their effect on anaemia risk in this region.

We fitted two Bayesian models; one logistic regression model to analyze the odds of anaemia across the entire population. The second model was fitted to assess the risk of anaemia due to malaria, Schistosomiasis, and helminthiasis in school-aged children (5-14 years). The second model was restricted to the school going ages because the schistosomiasis and helminthiasis prevalence data was predicted from surveys in this group (Kokaliaris et al., 2022), and did not fit well for the other ages and high extremely high uncertainty. The models were fitted in OpenBugs version 3.1.2 (Imperial College and Medical Research Council, London, UK), using the Markov Chain Monte Carlo (MCMC) simulation for parameter estimation. Spatial variation was treated as village-specific random effects, with latent observations of a spatial Gaussian process with a mean of zero and a covariance assuming an exponential variation function of distance between two villages (Diggle et al., 1998). We present the median posterior estimates from the Haemoglobin concentration Bayesian geostatistical models, while the anaemia categorical model's estimates were exponentiated to obtain odds ratios (OR). Each is accompanied by their 95% Bayesian Credible Intervals (BCI).

5.2.6. Bayesian model formulation

Let Y_{ij} be equal to 1 be the positive anaemia status and 0 otherwise of child *i* in village *j* at time *t, j=1,2,…385* and *t=1,2,…9* for the model fitted on the annual data. To estimate the odds of

being anaemic, we assume that Y_{ij} arises from a Bernoulli distribution, $Y_{ij} \sim Bernouli(p_{ij})$. We modelled the probability of being anaemic and other covariates (x) on the log scale of p_{ij}

$$logit(p_{ij}) = (\beta_0 + \beta_x + \phi_j) \qquad , k=1,2,...K$$

where N_{jt} the total person time contributed by persons in village j at time t (in person years, py), $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_K)^T$ are the regression coefficients (malaria prevalence and other predictors) and ϕ_j the village specific spatial effects.

For the model predicting the Hb concentration among 5–14-year-olds; the haemoglobin level of children in a village is assumed to come from normal distribution, that is $Y_{ij} \sim Normal(\mu_{ij}, \tau)$. Where the mean haemoglobin concentration of a child i, given the location j is given by:

$$\mu_{ij} = (\beta_0 + \beta_x X + \phi_j$$

In all the models, we assumed that $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots \phi_n)^T$ follows a Gaussian process, i.e. $\boldsymbol{\phi} \sim \text{MVN}(0, \sigma_1^2 R)$. Here, R is an exponential correlation matrix of the distance between villages, i.e. $R_{ij} = \exp(-\rho d_{kl})$ where d_{kl} is the distance between villages k and l, ρ is the rate of *correlation* decay with distance. The "effective range" refers to the shortest distance at which spatial correlation falls below 5% and is defined by $3/\rho$. The σ_1^2, is the variance of the spatial process. We specified non-informative normal prior distributions with mean zero and large variance for the regression coefficients $\beta_i \sim N(0, 10^3)$, an inverse gamma prior for σ_1^2, that is $\sigma_1^2 \sim IG(2.01, 1.01)$ and a Uniform prior distribution for ρ, that is $\rho \sim U(a, b)$, where a and b are chosen such that the effective range is within the maximum and minimum distances of the village's centroids.

The Bayesian models were then fitted using Markov Chain Monte Carlo (MCMC) simulation algorithm in OpenBugs version 3.1.2 (Imperial College and Medical Council, London, UK) to estimate model parameters (Gelfand & Smith, 1990). We initialized a single chain sampler for

50,000 iterations, discarding the first 5,000. Convergence was assessed by the Gelman-Rubin diagnostic (Gelman & Rubin, 1992) and was achieved after 5000 iterations.

5.2.7. Population attributable fraction of anaemia due to malaria, schistosomiasis and helminth infections

We estimated the Population attributable fraction (PAF) to anaemia due to either malaria prevalence or helminthiasis/schistosomiasis as shown below (Greenland et al., 2008);

$$PAF = \frac{P(OR - 1)}{P(OR - 1) + 1}$$

Where *P* is the posterior median prevalence of malaria or helminthiasis/schistosomiasis by age group, and *OR* the odds ratio (OR) for being anaemic in the same age group. The OR for malaria parasite presence or helminths prevalence were estimated by exponentiation of the median posterior estimates for anaemia obtained from the Bayesian geostatistical models above.

Ethical consideration

The study protocols for the KHDSS and parasite prevalence surveys received approval from the KEMRI's Scientific and Ethics Review Unit (SERU) as well as the CDC's Institutional Review Board (IRB). We obtained informed written consent from compound heads to conduct household surveys and only collected malaria prevalence data from those who provided individual consent, or in the case of minors, from their parents or guardians.

5.3. Results

5.3.1. Descriptive characteristics

Between 2007 and 2015, blood haemoglobin levels of 20,923 individuals were analysed. The average Hb concentration was 12.03 g/dL with a 95% Confidence Interval (CI) ranging between 11.98 and 12.07. The population anaemia prevalence was 42.3% (CI: 41.2-43.4%). Children aged 1-11 months had the highest anaemia rate at 67.1% (CI: 62.5-71.6%), followed by children aged 1-4 years at 61.0% (CI: 58.9-63.1%), and the elderly population at 47.1% (CI: 43.6-50.6%). Females had an anaemia prevalence of 45.3% (CI: 43.8-46.8%), notably higher than males. While individuals with malaria had a higher anaemia rate of 49.0% (CI: 47.2-50.8%) compared to those without malaria at 39.2% (CI: 37.8-40.5%). Those with clinical malaria had an even higher rate at 57.4% (CI: 53.5-61.3%), as opposed to those without clinical malaria at 41.3% (CI: 40.2-42.5%).

Malaria parasitaemia was relatively high at 36% (CI: 35.2-36.5%), and so was clinical malaria at 8.1% (CI: 7.7-8.4%). The prevalence of soil-transmitted helminths varied from 17% to 66%, averaging at 45.0% (CI: 44.9-45.2%), while schistosomiasis fluctuated between 1% and 29%, averaging 6.9% (CI: 6.8-7.0%).

Throughout this period, annual anaemia rates ranged between 20% and 34%, remaining relatively constant. A similar stable trend was observed for yearly Hb concentrations (Figure 5.1). Anaemia prevalence trends mirrored those of mean Hb concentrations within the same demographic groups. Children between the ages of 1-11 months and 1-4 years had the lowest mean Hb levels of 10.13% (CI: 9.93-10.34%) and 10.34% (CI: 10.26-10.41%) respectively. There were no significant differences in Hb levels or anaemia rates based on the risk associated with soil-transmitted helminths and Schistosomiasis (

Table 5.2).

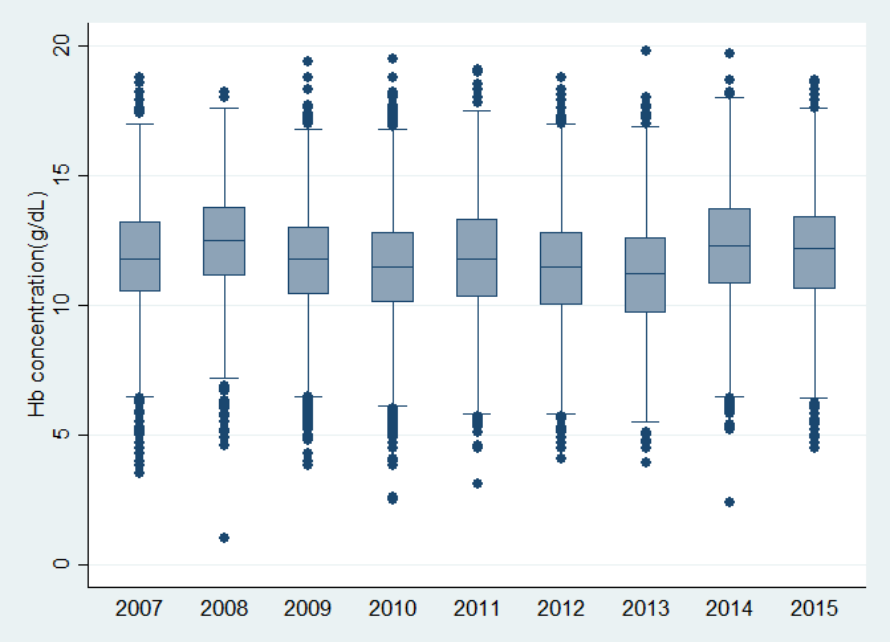


Figure 5.1: Box plot of Hb concentration over the years

Table 5.2: Population adjusted mean Hb concentration and anaemia prevalence among KHDSS residents between 2007 and 2015

Variable		Participants (%)	Mean HB level, g/dL (95% CI)	Anaemia prevalence (%) (95% CI)
Overall		20,923	12.03 (11.98-12.07)	42.3% (41.2%-43.4%)
Clinical Malaria	Negative	18,988(91.9)	12.10 (12.05-12.14)	41.3% (40.2%-42.5%)
	Positive	1,669 (8.1)	11.09 (10.89-11.28)	57.4% (53.5%-61.3%)
Parasite presence	Negative	13,216 (64.0)	12.25 (12.20-12.31)	39.2% (37.8%-40.5%)
	Positive	7,441 (36.0)	11.57 (11.49-11.64)	49.0% (47.2%-50.8%)
Schistosomiasis risk	<3.2%	6,714 (32.1)	12.07 (11.99-12.14)	42.2% (40.4%-43.9%)
	3.2-6.8%	6,955 (33.2)	12.00 (11.92-12.08)	42.7% (40.9%-44.6%)
	>6.8	7,254 (34.7)	12.02 (11.94-12.12)	41.9% (39.8%-44.0%)
Soil transmitted helminths risk	<39%	6,601 (31.5)	12.19 (12.11-12.26)	38.6% (36.8%-40.3%)
	39-50%	7,316 (35.0)	11.96 (11.88-12.05)	44.7% (42.7%-46.7%)
	>50%	7,006 (33.5)	11.96 (11.89-12.04)	42.8% (41.0%-44.6%)
Sex	Female	11,897 (55.6)	11.73 (11.67-11.78)	45.3% (43.8%-46.8%)
	Male	9,594 (44.4)	12.46 (12.39-12.54)	38.1% (36.6%-39.6%)
Age	1-11 months	971 (4.6)	10.13 (9.93-10.34)	67.1% (62.5%-71.6%)
	1-4 years	5,734 (27.4)	10.34 (10.26-10.41)	61.0% (58.9%-63.1%)
	5-14 years	5,809 (27.8)	11.93 (11.87-11.99)	36.5% (34.6%-38.3%)
	15-59 years	7,100 (33.9)	12.63 (12.56-12.71)	38.6% (36.9%-40.4%)
	60+ years	1,309 (6.3)	12.22 (12.07-12.36)	47.1% (43.6%-50.6%)
Year	2007	1,267 (6.1)	11.94 (11.75-12.13)	44.1% (40.0%-48.3%)
	2008	1,169 (5.6)	12.50 (12.33-12.67)	31.6% (27.7%-35.5%)
	2009	2,505 (12.0)	11.73 (11.62-11.85)	47.6% (44.4%-50.7%)
	2010	5,255 (25.1)	11.67 (11.59-11.75)	49.8% (47.8%-51.7%)
	2011	2,111 (10.1)	12.14 (12.00-12.28)	39.6% (36.2%-43.0%)
	2012	2,718 (13.0)	12.05 (11.93-12.17)	37.9% (35.0%-40.8%)
	2013	2,352 (11.2)	11.80 (11.67-11.93)	47.8% (44.7%-50.9%)
	2014	1,840 (8.8)	12.37 (12.24-12.50)	37.9% (35.1%-40.7%)
	2015	1,706 (8.2)	12.17 (12.05-12.30)	40.6% (37.6%-43.7%)
Elevation	>1200	3,646 (17.6)	12.07 (11.96-12.18)	40.1% (37.5%-42.8%)
	1201-1265	4,925 (23.7)	11.98 (11.88-12.08)	44.2% (41.6%-46.8%)
	1266-1300	3,508 (16.9)	12.02 (11.90-12.13)	41.7% (39.2%-44.3%)
	1301-1345	4,651(22.4)	11.98 (11.89-12.07)	43.9% (41.7%-46.1%)
	>1345	4,031 (19.4)	12.11 (12.02-12.21)	40.4% (38.1%-42.6%)
Region	Asembo	7,628 (36.5)	12.07 (11.99-12.15)	40.4% (38.4%-42.4%)
	Gem	6,421 (30.7)	12.03 (11.96-12.10)	43.0% (41.2%-44.8%)
	Karemo	6,874 (32.9)	11.97 (11.89-12.06)	43.8% (41.9%-45.6%)
SES	Poorest	6,126 (32.7)	11.98 (11.89-12.06)	43.5% (41.6%-45.4%)
	Poor	3,643 (19.5)	12.06 (11.95-12.16)	41.6% (39.1%-44.1%)
	Least Poor	8,956 (47.83)	12.12 (12.05-12.19)	40.9% (39.2%-42.5%)

Predicted risk of anaemia across all age groups

The logistic regression model for anaemia risk identified several factors that influenced the likelihood of anaemia in the population. Individuals with malaria parasitaemia had a higher

likelihood of anaemia (Odds ratio: OR = 1.82, 95% credible interval (CI): 1.69-1.96) compared to those without parasites. Interestingly, those diagnosed with clinical malaria had an even greater risk, with an OR = 2.03 (95% CI: 1.80-2.28), compared to individuals who only had the parasite or those who had the parasite but did not have fever.

Age was an important factor in determining the risk of anaemia. Children between the ages of 1 and 4 had the highest likelihood of developing anaemia, with an OR = 9.29 (95% CI: 7.49-11.51), followed by the 5-14 age bracket, with an OR = 5.49 (95% CI: 4.67-6.45). Compared to those under one year of age, these age groups were more vulnerable. On the other hand, the elderly, those aged 65 and above, showed the least susceptibility to anaemia.

Gender and socioeconomic factors were also important. Men had lower odds of anaemia (OR = 0.76, 95% CI: 0.71-0.82) compared to women. While, individuals at the top socioeconomic status were less likely to be anaemic (OR = 0.88, 95% CI: 0.82-0.96) compared to those from lower statuses. While a higher prevalence of conditions like soil-transmitted helminth and Schistosomiasis seemed to increase the anaemia risk, these variables were not statistically significant (

Table 5.3).

Table 5.3: Estimated posterior median odd ratios for anaemia risk and posterior median Hb concentrations for effects of malaria, schistosomiasis and soil transmitted helminth on Hb concentration

Variable		Odds ratios (95% BCI) (All age groups)	Median (95% BCI) (5-14 years)
*Clinical Malaria		**2.03 (1.80 - 2.28)**	**-0.52 (-0.67 to -0.36)**
Parasite presence		**1.82 (1.69 - 1.96)**	**-0.53 (-0.62 to -0.44)**
Male		**0.76 (0.71 - 0.82)**	-0.07 (-0.15 to 0.02)
Age	1-11 months	1	-
	1-4 years	**9.29 (7.49 - 11.51)**	-
	5-14 years	**5.49 (4.67 - 6.45)**	-
	15-59 years	0.94 (0.80 - 1.11)	-
	60+ years	0.78 (0.67 - 0.92)	-
Year	2007	1	**1**
	2008	**0.62 (0.50 - 0.77)**	**0.68 (0.42 to 0.94)**
	2009	**1.03 (0.85 - 1.24)**	-0.14 (-0.37 to 0.09)
	2010	**1.26 (1.08 - 1.48)**	**-0.41 (-0.60 to -0.21)**
	2011	**0.76 (0.59 - 0.99)**	-0.00 (-0.33 to 0.32)
	2012	0.82 (0.64 - 1.06)	-0.11 (-0.44 to 0.22)
	2013	1.21 (0.94 - 1.57)	**-0.44 (-0.77 to -0.12)**
	2014	0.85 (0.65 - 1.11)	-0.17 (-0.49 to 0.15)
	2015	0.99 (0.76 - 1.30)	-0.26 (-0.59 to 0.06)
Schistosomiasis risk		2.89 (0.43 - 20.42)	-1.17 (-3.39 to 0.98)
Soil transmitted helminths risk		1.62 (0.67 - 3.94)	-0.73 (-1.83 to 0.28)
Elevation	>1200	1	1
	1201-1265	0.91 (0.79 - 1.05)	-0.01 (-0.19 to 0.17)
	1266-1300	0.87 (0.74 - 1.02)	0.08 (-0.15 to 0.29)
	1301-1345	0.91 (0.77 - 1.07)	0.03 (-0.19 to 0.25)
	>1345	0.85 (0.71 - 1.02)	0.09 (-0.15 to 0.34)
Region	Asembo	1	1
	Gem	1.22 (0.71 - 1.02)	-0.19 (-0.61 to 0.23)
	Karemo	1.15 (0.85 - 1.66)	-0.12 (-0.57 to 0.41)
SES	Poorest	1	1
	Poor	0.94 (0.79 - 1.65)	-0.01 (-0.12 to 0.10)
	Least Poor	**0.88 (0.82 - 0.96)**	0.08 (-0.03 to 0.18
Range		8.09 (0.33 - 10.84)	11.24 (8.10 to 42.07)
$ Spatial variance		1.75 (0.27 - 6.87)	4.53 (1.22 to 16.81)

The estimates presented are the median of anaemia risk (OR) and 95% Bayesian credible intervals (BCI)

* Clinical malaria model estimates for other covariates not included but do not differ from the parasite presence model. Only provided here for comparison purposes

$ Minimum distance in kilometres at which spatial correlation is less than 5%

5.3.2. Predicted median haemoglobin concentration (4-5 years)

Similar variables from the logistic anaemia risk model were included in the model predicting factors affecting Hb concentration among school going children (

Table 5.3). Malaria parasitaemia was an important factor for predicting Hb concentration. Individuals with malaria parasitaemia had lower Hb concentration compared to those without malaria parasites [-0.53, (95%CI: -0.62 to -0.44)]. On the other hand, those with clinical malaria had lower median Hb concentrations compared to those without clinical malaria [-0.52, (95%CI: -0.67 to -0.36)]. There was not significant statistical differences on comparing the effects of parasite presence only versus clinical malaria even by age groups. Even though schistosomiasis prevalence showed a stronger influence on reduction of Hb concentration compared to soil transmitted helminths, this effect was not significant.

In this adjusted model, the year of survey was an important factor. There was a statistically significant downwards trend in the median Hb concentration by year of survey in comparison to the first survey, from 2007 [0.68 (0.42 to 0.94)] to 2015 [-0.26, (-0.59 to 0.06)]

Despite not being statistically important in this age group; being male was not protective, the highest socioeconomic status increased the median Hb concentration compared to the lower-level stratum. The minimum distance at which spatial correlation was below 5% was 11km in this age group, highlighting the importance of spatial variation in each group.

5.3.3. Population anaemia risk attributable to malaria parasites and helminths prevalence

The median Population Attributable Fraction (PAF) of anaemia due to the presence of malaria parasites was estimated to be 26% (95% CI: 23-29%), while for clinical malaria, the PAF was slightly higher at 31% (95% CI: 26%-36%). The impact of these factors varied across age groups. Among children aged 1-11 months, the attributable burden was the highest at 70%, followed by the 1–4-year age group at 48%. In contrast, the 15–59-year age group experienced the lowest attributable burden from these factors. While parasitaemia and clinical malaria accounted for a substantial proportion of the anaemia burden in the population Schistosomiasis and soil-transmitted helminths they were not statistically important.

5.4. Discussion

Based on cross-sectional population surveys, our study shows a persistently high prevalence of anaemia affecting more than two-fifths of the population, which is classified as a severe according to the WHO (World Health Organization, 2011). Our findings across various age groups reveal a sustained high prevalence of anaemia over time, predominantly affecting young children, women, and the elderly. Malaria parasitaemia, clinical malaria, age, gender, and socioeconomic status (which can be an indicator for malnutrition) were found to have a significant association with anaemia. However, there was no evident association between the prevalence of schistosomiasis/helminths and anaemia in this region.

The consistently high anaemia prevalence of over 42% over the 9-year span was on average above any available regional estimates (28% - 35%) (GBD Anaemia Collaborators, 2023; World Bank) also poses a major public health problem. Particularly alarming is the anaemia rate among children. Specifically, those between 1-11 months and 1-4 years, with prevalence rates of 67% and 40% respectively. These figures starkly contrast the 40% and 26% results from a concurrent national malaria survey (Ngesa & Mwambi, 2014) carried out during the same period. Research suggests that children under the age of 5 years are more prone to anaemia when exposed to malaria. This heightened vulnerability can be attributed to their still-evolving immune systems, which leaves them more susceptible to infections like malaria, subsequently intensifying anaemia (Alusala et al., 2008; Marsh & Snow, 1999; Menendez et al., 2000). In contrast, adults and the elderly population might experience increased anaemia risk due to age-related declines in overall health, combined with exposure to other risk factors like malaria and HIV/AIDS, which are also drivers of anaemia (Safiri et al., 2021).

Malaria was the most important determinant of Hb concentration and anaemia risk. Malaria, is known to cause haemolysis and hinders the normal development of red blood cells (Menendez

et al., 2000; White, 2018), resulting in anaemia. The significant reduction in median Hb concentration associated with malaria and the substantially heightened risk of anaemia in individuals exhibiting clinical malaria symptoms compared to those with parasitaemia only underlines its impact on haemoglobin concentration. This strong association between malaria parasite prevalence and anaemia bolsters existing research highlighting its contribution to the anaemia, especially in malaria-prone areas (Ejigu et al., 2018; Kassebaum, 2016; Soares Magalhães & Clements, 2011; Sumbele et al., 2013). Age-specific studies in the region have also evidenced this connection, particularly among adolescents and expectant mothers (Desai et al., 2005; Foote et al., 2013; Koukounari et al., 2008; Ouma et al., 2007). The fact that nearly half of those with malaria parasitaemia suffer from anaemia, as opposed to two out of five without the parasite, further emphasizes the association between malaria and anaemia. It underlines the importance of continued robust malaria control measures in regions with high transmission rates (Noor, Kirui, et al., 2009), not just in this area but to a larger extend the whole lake region of western Kenya and others regions with a similar malaria profiles.

Surprisingly, our study did not observe any significant association between anaemia and the prevalence of soil-transmitted helminths or Schistosomiasis. This stands in contrast to several studies that have previously linked these infections to anaemia (Douglas et al., 2012; Koukounari et al., 2008; Ouma et al., 2007; Sumbele et al., 2013; White, 2018). Despite this lack of association in our study, their prevalence is still of public health concern. It is possible that the interplay between malaria, STH, and helminths is complex, with malaria potentially overshadowing or altering the impact of helminths. Additionally, the two might have intricate interactions that necessitate more in-depth investigations (Brooker et al., 2007; Mwangi et al., 2006; Nacher, 2011). The absence of a clear link could also stem from our study's limited geographical scope, possibly creating a consistent effect of Schistosomiasis and helminths prevalence. Moreover, it is plausible that the prevalence estimates for STH and helminths, based on school-age children's surveys, might not truly reflect the broader community's situation.

Gender, as seen in our study plays a significant role in anaemia prevalence, with females exhibiting a higher likelihood and prevalence than males. Previous research (Kassebaum, 2016; Safiri et al., 2021; Stevens et al., 2013) supports this observation, linking it to unique challenges women face, especially in low-income countries. These challenges include menstrual blood loss, pregnancy complications, childbirth, and inadequate diets, which contribute to reduced iron levels in their bodies (Derman & Patted, 2023), Consequently, women are more burdened by anaemia than men.

Higher socio-economic status was a protective factor against anaemia. Socio-economic status has been shown to be a key indicator of malnutrition in low-income countries (Müller & Krawinkel, 2005; World Health Organization, 2017b; Yang et al., 2018). Those with at lower socio-economic status often face poor food quality, food insecurity, and inadequate nutrition (Safiri et al., 2021). It could also be that benefits of a higher socio-economic status, such as better nutrition, healthcare access, and living conditions, can mitigate the effects of malaria and other factors causing anaemia, especially in high-risk age groups.

Anaemia prevalence across the nine-year period was stable, which could be indicative of existing interventions including folic iron supplementation (Division of National Malaria Programme (DNMP) & ICF, 2021; Ministry of Health, 2013), not achieving the intended objectives. However, many of these interventions target primarily pregnant women and younger children.

Anaemia prevalence's spatial correlations hint at common environmental, social, or healthcare influences. This spatial distribution could also highlight localized interactions, such as soil-transmitted helminth prevalence, that might be overlooked in broader analysis. Recognizing these specific patterns can inform targeted interventions and community strategies.

Our study lacked data on dietary habits, genetic factors, HIV/AIDS, and other regional infectious diseases. Additionally, the use of aggregated data for STH and schistosomiasis introduced significant uncertainty. Assuming that prevalence data, modelled from surveys of

school-aged children, is representative may have introduced bias. These limitations should be factored in when interpreting our results. Nonetheless, the use of village-aggregated data enabled us to assess the anaemia burden across all age groups, distinguishing our research from numerous other studies.

5.5. **Conclusion**

The high prevalence of anaemia in this region of western Kenya, especially among specific vulnerable groups, necessitates urgent and targeted interventions. Multiple important predictors, including age, gender, socio-economic status, and malaria status, indicate a complex interplay that reinforces the notion that anaemia is not solely a one direction problem but rather a multifaceted public health challenge that requires an integrated approach. The importance of malaria control, gender-specific strategies, age-targeted care, and the consideration of underlying socio-economic factors in designing comprehensive public health interventions for anaemia control should be considered.

Declarations

Ethical consideration

The study protocol for the KHDSS household and parasitaemia surveys has been approved by the Kenya medical research institute (#1801) and Centers for Disease Control and Prevention (#3308) institutional review boards. Informed written consent was obtained from the heads of compounds in conducting the household interviews and additional consent was sought from individuals or the caretakers of those under 18 years who participated in the health facility surveillance.

Data availability

Data were obtained with permission of the KEMRI-CGHR HDSS and Malaria branch steering committee. Any data requests may be sent to the respective steering committees, through Dr. Simon Kariuki (Skariuki@kemricdc.org) for malaria parasitaemia surveys or Dr. Stephen Munga (Smunga@kemri.org) for the KHDSS data.

Competing interests

The authors declare no competing interests.

Funding

SK was supported by the European Research Council (ERC) of Penelope Vounatsou grant number 323180 and the Swiss National Foundation (SNF) program for Research on Global Issues for Development (R4D) project number IZ01Z0-147286 as part of his PhD work. The funders did not have any role in the study or its outcome.

Authors' contributions

SK and PV conceived the study; SK analysed the data and drafted the initial manuscript; PV supported data analysis; SK, MD, KO, LS, JG, AO, FO, AS, SK and MH designed the initial surveillance system, participated in data collection, processing and management. All authors gave inputs in revision of the first drafts, read and approved the final manuscript.

Acknowledgement

We thank the HDSS community in Siaya County for the many years of support and participation in surveillance activities; the HDSS and Malaria branch staff of the KEMRI/CDC collaboration both past and present for assisting in managing the platforms, data collection and management. This article is published with the permission of the director Kenya Medical Research Institute.

5.6. Appendix

5.6.1. Posterior estimates for effects of malaria, Schistosomiasis and soil transmitted helminths prevalence and other factors on Hb concentration

Covariate		1-11 mon Median (95% BCI)	1-4 yrs. Median (95% BCI)	5-14 yrs. Median (95% BCI)	15-59 yrs. Median (95% BCI)	60 plus Median (95% BCI)	Overall* Median (95% BCI)
Clinical malaria**		-1.45 (-1.89 to -1.01)	-0.99 (-1.13 to -0.84)	-0.52 (-0.67 to -0.36)	-0.19 (-0.44 to 0.06)	-0.33 (-0.84 to 0.19)	-0.70 (-0.81 to -0.60)
Parasite presence		-1.52 (-1.79 to -1.23)	-0.92 (-1.02 to -0.82)	-0.53 (-0.62 to -0.44)	-0.19 (-0.31 to -0.08)	-0.19 (-0.53 to 0.15)	-0.55 (-0.61 to -0.49)
Sex (Male)		-0.30 (-0.54 to -0.07)	-0.10 (-0.19 to 0.00)	-0.07 (-0.15 to 0.02)	2.05 (1.95 to 2.15)	0.81 (0.56 to 1.07)	0.64 (0.59 to 0.70)
Year	2007	1	1	1	1	1	1
	2008	0.39 (-0.35 to 1.14)	0.22 (-0.15 to 0.59)	0.68 (0.42 to 0.94)	0.57 (0.31 to 0.83)	1.01 (0.37 to 1.64)	0.52 (0.36 to 0.68)
	2009	0.46 (-0.19 to 1.10)	-0.50 (-0.80 to -0.20)	-0.14 (-0.37 to 0.09)	0.01 (-0.23 to 0.24)	0.56 (0.00 to 1.13)	-0.08 (-0.23 to 0.07)
	2010	-0.44 (-1.00 to 0.12)	-0.58 (-0.84 to -0.32)	-0.41 (-0.60 to -0.21)	0.005 (-0.20 to 0.21)	0.32 (-0.18 to 0.81)	-0.25 (-0.37 to -0.13)
	2011	0.16 (-0.55 to 0.89)	-0.19 (-0.60 to 0.18)	0.00 (-0.33 to 0.32)	0.51 (0.19 to 0.85)	0.99 (0.26 to 1.72)	0.23 (0.01 to 0.44)
	2012	0.53 (-0.17 to 1.23)	-0.28 (-0.67 to 0.09)	-0.11 (-0.44 to 0.22)	0.37 (0.05 to 0.70)	1.42 (0.70 to 2.15)	0.21 (0.00 to 0.42)
	2013	-0.07 (-0.80 to 0.66)	-0.86 (-1.26 to -0.49)	-0.44 (-0.77 to -0.12)	0.25 (-0.07 to 0.59)	0.59 (-0.17 to 1.35)	-0.23 (-0.45 to -0.02)
	2014	0.35 (-0.46 to 1.17)	-0.42 (-0.86 to 0.01)	-0.17 (-0.49 to 0.15)	0.00 (0.28 to 0.92)	0.00 (0.13 to 1.49)	0.25 (0.03 to 0.47)
	2015	0.11 (-0.99 to 1.21)	-0.69 (-1.14 to -0.24)	-0.26 (-0.59 to 0.06)	0.27 (-0.05 to 0.60)	0.65 (-0.06 to 1.38)	0.03 (-0.19 to 0.25)
Schistosoma risk		-0.74 (-4.43 to 3.10)	-1.55 (-4.01 to 0.98)	-1.17 (-3.39 to 0.98)	0.16 (-1.73 to 2.04)	2.69 (-0.78 to 6.06)	-1.71 (-3.50 to -0.04)
Helminths risk		0.83 (-1.47 to 3.07)	-1.44 (-2.81 to -0.27)	-0.73 (-1.83 to 0.28)	-0.68 (-1.68 to 0.40)	1.15 (-1.05 to 3.28)	-0.61 (-1.40 to 0.12)
Elevation	>1200	1	1	1	1	1	1
	1201-1265	0.25 (-0.18 to 0.69)	0.23 (0.03 to 0.43)	-0.01 (-0.19 to 0.17)	-0.02 (-0.21 to 0.16)	0.11 (-0.32 to 0.53)	0.07 (-0.04 to 0.19)
	1266-1300	0.11 (-0.39 to 0.62)	0.13 (-0.10 to 0.36)	0.08 (-0.15 to 0.29)	0.09 (-0.13 to 0.31)	0.14 (-0.35 to 0.64)	0.07 (-0.07 to 0.21)
	1301-1345	0.09 (-0.40 to 0.59)	0.27 (0.04 to 0.52)	0.03 (-0.19 to 0.25)	0.08 (-0.14 to 0.30)	0.28 (-0.24 to 0.78)	0.12 (-0.02 to 0.27)
	>1345	0.09 (-0.45 to 0.64)	0.26 (-0.01 to 0.52)	0.09 (-0.15 to 0.34)	0.13 (-0.10 to 0.37)	0.43 (-0.12 to 0.98)	0.11 (-0.04 to 0.28)
Region	Asembo	1	1	1	1	1	1
	Gem	-0.69 (-1.18 to -0.19)	-0.14 (-0.63 to 0.38)	-0.19 (-0.61 to 0.23)	-0.14 (-0.44 to 0.15)	-0.25 (-0.71 to 0.23)	-0.08 (-0.59 to 0.37)
	Karemo	-0.51 (-1.01 to -0.02)	-0.04 (-0.57 to 0.48)	-0.12 (-0.57 to 0.41)	-0.16 (-0.47 to 0.25)	-0.30 (-0.76 to 0.15)	-0.01 (-0.62 to 0.48)
SES	Poorest	1	1	1	1	1	1
	Poor	0.05 (-0.23 to 0.33)	0.14 (0.02 to 0.26)	-0.01 (-0.12 to 0.10)	0.02 (-0.10 to 0.14)	0.26 (-0.06 to 0.57)	0.06 (-0.01 to 0.13)
	Least Poor	-0.0758 (-0.37 to 0.21)	0.07 (-0.05 to 0.19)	0.08 (-0.03 to 0.18)	0.15 (0.04 to 0.27)	0.21 (-0.07 to 0.49)	0.11 (0.04 to 0.18)
Range		24.13 (8.37 to 92.48)	11.02 (8.10 to 36.25)	11.24 (8.10 to 42.07)	20.39 (8.28 to 90.43)	28.94 (8.54 to 94.49)	11.00 (8.10 to 39.99)
Spatial variance[$]		0.70 (0.18 to 5.51)	6.58 (1.86 to 23.03)	4.53 (1.22 to 16.81)	1.00 (0.24 to 7.01)	0.45 (0.16 to 2.11)	13.56 (7.36 to 47.76)

Estimates presented here are median of Hb concentrations and 95% Bayesian credible intervals (BCI) adjusted for other covariates

[*] Age-adjusted

[**] Clinical malaria model estimates for other covariates not included but do not differ from the parasite presence model. Only provided here for comparison purposes

[$] Minimum distance in kilometres at which spatial correlation is less than 5%

6. Chapter 6: General discussion

Using data from the KEMRI Health and Demographic Surveillance System (KHDSS) located in western Kenya, we investigated the spatio-temporal dynamics of malaria incidence and parasite prevalence as measures of transmission and their association with both all-cause and malaria-specific mortality across all age categories. We also explored the anaemia-related burden due to malaria in the same area.

This thesis comprises of four objectives addressed in chapters 2-5, where detailed methodology, results, discussions and conclusions have been provided. Based on these findings and prevailing knowledge, we provide an overall discussion of all the chapters. Specifically, we relate the findings and come up with, significant key findings, contribution to epidemiological methods, malaria epidemiology, future outlook and extension of this work.

6.1. Malaria epidemiology

6.1.1. High malaria burden and the effect of climatic/environmental factors

In the KHDSS, the epidemiology of malaria paints a concerning picture, especially for children under the age of five. With incidence rates ranging between 450 and 775 cases per 1,000 person-years of observation (pyo), the statistics are staggeringly higher than the regional average of less than 100 cases pyo during the same period. This vast difference underscores an immediate need for specific and localized interventions for malaria control.

Given that these high incidences are not just temporary spikes, the stability of such high rates over an extended period renders the situation even more pressing. Analysing the transmission patterns, a distinct cyclical nature becomes evident, with peaks consistently observed between March-July and September-January. This seasonality aligns with the short and long rainy seasons. Recognizing this pattern provides a clear window of opportunity where preventive measures can be deployed more intensively, potentially curbing the rates of transmission.

Individual climatic conditions, such as temperature and rainfall, do not seem to have a statistically important relationship with malaria transmission. However, considering their cumulative effect through the lens of the Enhanced Vegetation Index (EVI), the narrative shifted. There was a significant association between denser vegetation, as represented by higher EVI values, and the incidence of malaria. This suggests that while individual climatic factors might not stand out on their own, their combined effect creates a conducive environment that drives malaria transmission.

Such insights highlight the importance of understanding the intricacies of malaria transmission even on a small scale. Yet, it is essential to note that this thesis identified certain shortcomings in the data, which led to subsequent analyses that were not covered here, but considered in our discussions. The subsequent studies on a KHDSS subpopulation that is monitored more intensely, further shed light on the critical role of climate variability in malaria transmission (Beloconi et al., 2023; Nyawanda et al., 2023). Their findings revealed that climate factors have a higher significant influence on malaria transmission than interventions like mosquito nets and indoor residual spraying. Echoing these findings, others studies have also established that adverse climate parameters further challenge malaria control efforts, amplifying transmission conditions through favourable temperature and rainfall environments (Lubinda et al., 2021).

6.1.2. Malaria mortality

Apart from environmental, climatic and other known interventions like bed nets, indoor residual spraying and treatment, malaria mortality is directly impacted in a broader context by the community's social and economic landscape. This can encompass various factors including, the political environment, concurrent health threats like epidemics (for instance, COVID-19) (Weiss et al., 2021), civil unrest (Feikin et al., 2010), and natural calamities. One noticeable observation from our data is the prominent influence of external shock events, such as post-

election violence in 2008, on mortality rates. The aftermath of such events clearly demonstrates that while the transmission rates of malaria (incidence and prevalence) did not skyrocket immediately, the mortality rates did. The disparity between these rates during such times highlights the vulnerability of health systems under societal stress. Interestingly, while death rates saw a steep decline post-2008, the incidence and prevalence exhibited a more gradual reduction, indicating a lingering effect of the violence on overall malaria transmission.

In light of these findings, we postulate that incidence data from functional health information registration systems at health facilities (as measured from select health facilities in the KHDSS) offer a more accurate picture of the malaria-mortality relationship than intermittent prevalence surveys. The reasoning is twofold: not only does a well-functioning health reporting system provide more reliable estimates, it also plays a pivotal role in ensuring a comprehensive capture of cases, facilitating proper diagnosis and treatment.

6.1.3. Measures of malaria transmission and their relationship with mortality

Comparing different measures of transmission, we noted that incidence is a better predictor of mortality than EIR or PP in the short term (Khagayi et al., 2017; Khagayi et al., 2019). In the realm of malaria transmission measures and their impact on mortality, both incidence and parasite prevalence (PP) hold significant value, albeit serving distinct purposes.

Incidence delineates the number of new cases emerging within a population over a specific timeframe. It is particularly adept at capturing short-term impacts, far more so than Entomological Inoculation Rate (EIR) and PP. This immediacy allows incidence data to promptly spotlight changes in malaria transmission and underscores its sensitivity to abrupt disruptions to the health system. Such disruptions can range from economic downturns and civil unrest to sudden epidemics. On the other hand, despite its potential to capture immediate factors pivotal for monitoring malaria burdens, such as seasonal shifts, its sensitivity can be a limitation.

For instance, the study revealed that incidence was not a robust predictor of overall population mortality. However, the research encountered challenges in gauging health-seeking behaviours accurately, which can considerably influence incidence measurements.

Contrastingly, PP offers a snapshot of the proportion of a population infected with the malaria parasite at a given moment. Its strength lies in its capacity to provide a consistent, long-term view of malaria transmission. For instance, five-year aggregated PP, as a transmission parameter, provides a comprehensive picture of the region's malaria status since it is not affected by seasonality, civil unrest, or other instantaneous shocks to the health system. Within the study's framework, PP, particularly the slide positivity rate, outperformed incidence as a predictor of mortality. This can be attributed to the observation that symptomatic adults, who are often captured by prevalence data, do not consistently seek medical care as compared to children (Bigogo et al., 2010). Yet, just like incidence, PP is not devoid of limitations. It may fall short in capturing on-the-spot factors that influence malaria transmission, especially when juxtaposed with the dynamic data that incidence offers.

The subtle difference between these two measures underscores the importance of a multi-faceted approach in studying malaria transmission. For a holistic view, there is a pressing need for regular household surveys to integrate components addressing health-seeking behaviours, ensuring robust data collection. Embracing the strengths and understanding the limitations of both incidence and prevalence can usher in more targeted and effective interventions, particularly in regions grappling with external challenges.

6.1.4. Malaria in the 5-14 age group

Although older children displayed lower overall and malaria-specific mortality rates, this should not lead to complacency. Interestingly, despite a lower clinical malaria incidence and slide positivity rate (SPR) in these children (Khagayi et al., 2017), they exhibit a high

prevalence of parasitaemia (Khagayi et al., 2019). This high parasitaemic state underscores their potential role as reservoirs, threatening not just their health but the wider community through the potential for disease transmission (Beloconi et al., 2023). Recent research emphasizes the role of older, asymptomatic children as significant contributors to the transmission dynamics of malaria in areas with high disease prevalence (Andagalu et al., 2023; Rek et al., 2022). Moreover, climate change, has begun to play an increasingly influential role in determining the disease patterns and impacts among older children (Lubinda et al., 2021).

Mortality rates in these older children were the lowest compared to other age groups, however, the overall contribution of malaria as a cause of death in the group was quite high. Persistent parasitaemia, although subtle, exposes these children to continuous bouts of clinical malaria, subsequently increasing the risk of severe complications or even death. This hypothesis gains strength when considering that home-based measurements of clinical malaria, which provide an almost immediate insight into a child's health status, closely mirrored the incidence data from hospitals. Both point towards a strong correlation between clinical malaria and mortality in older children. However, one of the most pressing concerns in addressing malaria in this age group is the potential delay in seeking treatment. There's a prevailing notion that older children, due to their perceived resilience compared to infants, might not be prioritized when it comes to seeking immediate medical care (Bigogo et al., 2010). This potential delay or negligence can lead to exacerbated health complications, turning an otherwise treatable condition into a life-threatening one. Another area of concern is the evolving mortality patterns. With a spotlighted potential shift in mortality from infants to older children (Desai et al., 2014), suggesting that current malaria interventions might not be benefiting the latter as much. This becomes even more concerning when paired with findings about the high rates of treatment failures using artemether-lumefantrine among older children (Andagalu et al., 2023).

Understanding malaria in older children is therefore of concern. While they might initially seem less vulnerable than infants, various biological, societal, and environmental factors may compound their risk, and increase the reservoir for infection to more vulnerable categories like pregnant women and younger children.

6.1.5. Malaria anaemia relationship

Malaria parasitaemia and clinical malaria were identified as significant determinants of anaemia. This intricate relationship, suggesting that malaria infection does act as a direct causative agent of anaemia. This presents a 'double jeopardy' phenomenon in this region, because extremely high levels of malaria not only exacerbate anaemia but also accentuate the effects of certain genetic conditions that lead to anaemia. Such conditions have been highlighted in other studies (Foote et al., 2013), indicating a persistent problem in the region. Complicating this relationship further are other potential determinants of anaemia, such as HIV/AIDS, malnutrition, other infectious diseases, and inherent genetic conditions (GBD Anaemia Collaborators, 2023). This paints a multi-faceted picture of anaemia's aetiology in the region, emphasizing the need for a more comprehensive understanding.

Considering the intricacies of anaemia in this region, further studies are needed. These studies should not only focus on the direct relationship between malaria and anaemia but also aim to elucidate how multiple conditions and factors converge to influence anaemia's prevalence and severity. While the immediate health implications of this intricate relationship are evident, the long-term effects on the population remain less clear. Does persistent malaria-induced anaemia lead to other chronic health conditions? Does it decrease life expectancy, reduce the quality of life, or impair cognitive development in children? Addressing these questions is imperative to implementing effective public health interventions.

6.2. Contribution of the thesis

This research on the geo-statistical modelling of malaria mortality and its association with anaemia in the KHDSS, Kenya, is multi-faceted in its significance. It uncovers the complexities of malaria transmission in the region, emphasizing that a comprehensive approach, which considers seasonality, climate, socio-economic conditions, and geography, is vital. Targeted interventions in high-risk zones during peak transmission periods may yield better results than broad, generalized strategies.

The study underscores the need for enhanced malaria control measures. By highlighting the strong correlation between malaria transmission intensity and mortality, it offers a clear roadmap for health officials. With resources now devolved to local regions after the devolution of health policy (Masaba et al., 2020), there is an opportunity for better resource allocation, especially in regions like KHDSS that bear a substantial malaria burden as documented in this study and also shown by the Division of National Malaria Programme (DNMP) (Division of National Malaria Programme (DNMP) & ICF, 2021). Given that older children serve as a significant parasite reservoir in this region, it is essential that resources be optimized for comprehensive malaria prevention and treatment – from universal bed-net coverage to prompt provision of antimalarial drugs at health centres.

Moreover, the research draws attention to the secondary impacts of malaria, notably anaemia, across all age groups. By establishing a link between malaria and the consistent prevalence of anaemia, it calls for integrated health strategies that address both the primary disease and its cascading health effects. It is a clarion call to widen the scope of interventions beyond just young children and pregnant women, given the pervasive anaemia burden.

The statistical methodologies used in this study, such as Bayesian hierarchical negative binomial models and spatio-temporal survival models, further enrich the global scientific community's toolkit. While this study is domiciled in the KHDSS, its methodologies and

findings can serve as a template for similar research in other malaria-affected areas, underscoring the importance of understanding regional disparities for global malaria eradication initiatives.

In summary, this research is pivotal for reshaping malaria control measures, both within KHDSS and beyond. By presenting a detailed overview of malaria's patterns, mortality, and associated complications, it equips a range of stakeholders—from local health experts to global entities—to tackle malaria with enhanced precision and efficacy.

6.3. Limitations and challenges

Potential challenges and limitations during our research would have been due to the following:

Time period covered: Our research spanned from 2007 to 2015, with some datasets extending only to 2012. This timeframe may not fully capture the evolving dynamics of malaria's impact. Moreover, changes in malaria incidence or management post-2015 might not be reflected in our findings. However, this work has led to further refined work spanning an extra five years.

Malaria-anaemia association: While our study showed a strong association of malaria and anaemia prevalence, it is crucial to recognize that anaemia can arise from various causes. Although we controlled for Schistosomiasis/helminthiasis prevalence, other unknown factors whose data was not collected i.e., pregnancy status, dietary habits, genetic conditions, HIV/AIDS and other underlying health conditions that are prevalent in this region could also contribute to anaemia's high prevalence.

Health-seeking behaviour adjustments: The limitations around adjusting for health-seeking behaviour (one-time survey in the KHDSS) might have some bearing on the results. The behaviour might have been modified over the years or influenced by changes in policy. However, subsequent in-depth analysis of a smaller sub-population that was actively monitored for illnesses provides a more nuanced understanding and rectify the shortcomings of this broader study.

Other external factors: Even though our study incorporated numerous variables, it might not have fully considered the influence of factors like local health campaigns, presence of other endemic diseases, shifts in micro-policy, emerging diseases, outbreaks, or alterations in healthcare practices. Any of these elements could potentially affect malaria incidence and mortality.

Recognizing and understanding these challenges and limitations is essential for refining future research and ensure that the study findings are interpreted within the appropriate context.

6.4. Extension and future research

Given the findings, significance, and limitations of the study, the following would be valuable extensions of our work for future research:

1. **Extended temporal analysis and model fine-tuning:** An updated and extension of the temporal scope of the study to include data from post-2015 to the present to assess any recent trends or changes in the malaria landscape. This would take into consideration other recent events of public health importance, including the COVID-19 pandemic (Weiss et al., 2021), policy changes in other diseases like test and treat for HIV which has been implemented in Kenya among others. The effect of these events could inform more on ways of dealing with similar public health shocks in future. In line with this, better fine-tuned models should be developed to address potential issues in Bayesian modelling assumptions and refine models based on improved statistical methodologies or algorithms, better-informed priors, updated data, and potential confounders not included in our research would be useful in improving accuracy and predictive power. Such models have been implemented (Beloconi et al., 2023; Nyawanda et al., 2023) and other are being developed in line with our findings with an aim of developing localised early warning systems.
2. **Comprehensive anaemia-malaria studies on the whole population:** Due to the notable anaemia prevalence, research across all age groups, focusing on its varied causes in the region, not limited to malaria and Schistosomiasis/helminths but extending to iron deficiency, HIV/AIDS, malnutrition, and genetic factors, is essential.
3. **Health systems research and implementation studies:** Given the understanding of malaria's burden in the KHDSS region, implementation studies on the effectiveness of various interventions, tailored to the local context would be the logical next step. In addition, understanding how health systems' responsiveness, efficiency, and access can influence malaria incidence, prevalence and mortality.

4. **Climate change and malaria:** Given the observed relationships with climatic variables, studying how climate change might impact future malaria patterns in the region would be pertinent. Emerging evidence (Lubinda et al., 2021; Mafwele & Lee, 2022; Ryan et al., 2020), suggests that climate change will result in increased risk of malaria transmission. Making this an important research theme to explore.
5. **Malaria mortality determination:** With the important role of mortality data in tracking malaria control progress, VA in general and as a tool for malaria-mortality diagnosis needs to be improved despite the noted shortcomings. Minimally invasive tissue sampling (MITS) - a simplified post-mortem examination technique - that has shown to be an adequate approach for cause of death investigation in low-resource settings (Bassat et al., 2013; Rakislova et al., 2019), could potentially provide data to help quantify the relative contribution of malaria. A thorough investigation of cause of death using MITS techniques could potentially provide data to help quantify the relative contribution of the role of malaria to mortality. The data should subsequently be utilized to improve VA cause of death determination algorithms and hence its diagnostic ability.

7. Bibliography

Adair, T. (2021). Who dies where? Estimating the percentage of deaths that occur at home. *BMJ Global Health, 6*(9), e006766. https://doi.org/10.1136/bmjgh-2021-006766

Adazu, K., Lindblade, K. A., Rosen, D. H., Odhiambo, F., Ofware, P., Kwach, J., . . . Slutsker, L. (2005). Health and Demographic Surveillance in Rural Western Kenya: A Platform for Evaluating Interventions to Reduce Morbidity and Mortality from Infectious Diseases. *The American Journal of Tropical Medicine and Hygiene, 73*(6), 1151-1158.

Afrane, Y. A., Zhou, G., Githeko, A. K., & Yan, G. (2014). Clinical malaria case definition and malaria attributable fraction in the highlands of western Kenya. *Malar J, 13*, 405. https://doi.org/10.1186/1475-2875-13-405

Alegana, V. A., Atkinson, P. M., Lourenço, C., Ruktanonchai, N. W., Bosco, C., Erbach-Schoenberg, E. z., . . . Tatem, A. J. (2016). Advances in mapping malaria for elimination: fine resolution modelling of Plasmodium falciparum incidence. *Scientific Reports, 6*, 29628. https://doi.org/10.1038/srep29628

Alegana, V. A., Khazenzi, C., Akech, S. O., & Snow, R. W. (2020). Estimating hospital catchments from in-patient admission records: a spatial statistical approach applied to malaria. *Scientific Reports, 10*(1), 1324. https://doi.org/10.1038/s41598-020-58284-0

Alemu, A., Shiferaw, Y., Addis, Z., Mathewos, B., & Birhan, W. (2013). Effect of malaria on HIV/AIDS transmission and progression. *Parasites & Vectors, 6*, 18. https://doi.org/10.1186/1756-3305-6-18

Alusala, D. N., Estambale, B. B. A., Magnussen, P., Friis, H., Luoba, A. I., & Mwaniki, D. (2008). Predictors of serum ferritin and haemoglobin during pregnancy, in a malaria-endemic area of western Kenya. *Annals of Tropical Medicine and Parasitology, 102*(4), 297-308. https://doi.org/10.1179/136485908X278874

Amek, N. (2013). *Bayesian spatio-temporal modelling of the relationship between mortality and malaria transmission in rural western Kenya* [PhD, University_of_Basel]. edoc.unibas.ch.

Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J., Laserson, K. F., . . . Vounatsou, P. (2011). Spatio-temporal modeling of sparse geostatistical malaria sporozoite rate data using a zero inflated binomial model. *Spatial and Spatio-Temporal Epidemiology, 2*(4), 283-290. https://doi.org/10.1016/j.sste.2011.08.001

Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J. E., Odhiambo, F., . . . Vounatsou, P. (2012). Spatial and temporal dynamics of malaria transmission in rural Western Kenya. *Parasites & Vectors, 5*(1), 86. https://doi.org/10.1186/1756-3305-5-86

Amek, N., Van Eijk, A., Lindblade, K. A., Hamel, M., Bayoh, N., Gimnig, J., . . . Vounatsou, P. (2018). Infant and child mortality in relation to malaria transmission in KEMRI/CDC HDSS, Western Kenya: validation of verbal autopsy. *Malaria Journal, 17*, 37. https://doi.org/10.1186/s12936-018-2184-x

Amek, N., Vounatsou, P., Obonyo, B., Hamel, M., Odhiambo, F., Slutsker, L., & Laserson, K. (2015). Using health and demographic surveillance system (HDSS) data to analyze geographical distribution of socio-economic status; an experience from KEMRI/CDC HDSS. *Acta Trop, 144*, 24-30. https://doi.org/10.1016/j.actatropica.2015.01.006

Amek, N. O., Odhiambo, F. O., Khagayi, S., Moige, H., Orwa, G., Hamel, M. J., . . . Laserson, K. F. (2014). Childhood cause-specific mortality in rural Western Kenya: application of the InterVA-4 model. *Glob Health Action, 7*, 25581. https://doi.org/10.3402/gha.v7.25581

Amir, A., Cheong, F. W., de Silva, J. R., Liew, J. W. K., & Lau, Y. L. (2018). Plasmodium knowlesi malaria: current research perspectives. *Infect Drug Resist, 11*, 1145-1155. https://doi.org/10.2147/idr.S148664

Anand, I. S. (2008). Anemia and Chronic Heart Failure: Implications and Treatment Options. *Journal of the American College of Cardiology, 52*(7), 501-511. https://doi.org/10.1016/j.jacc.2008.04.044

Andagalu, B., Watson, O. J., Onyango, I., Opot, B., Okoth, R., Chemwor, G., . . . Kamau, E. (2023). Malaria Transmission Dynamics in a High-Transmission Setting of Western Kenya and the Inadequate Treatment Response to Artemether-Lumefantrine in an Asymptomatic Population. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, 76*(4), 704-712. https://doi.org/10.1093/cid/ciac527

Bailo, P., Gibelli, F., Ricci, G., & Sirignano, A. (2022). Verbal Autopsy as a Tool for Defining Causes of Death in Specific Healthcare Contexts: Study of Applicability through a Traditional Literature Review. *Int J Environ Res Public Health, 19*(18). https://doi.org/10.3390/ijerph191811749

Balarajan, Y., Ramakrishnan, U., Özaltin, E., Shankar, A. H., & Subramanian, S. (2011). Anaemia in low-income and middle-income countries. *The Lancet, 378*(9809), 2123-2135. https://doi.org/10.1016/S0140-6736(10)62304-5

Bartoloni, A., & Zammarchi, L. (2012). Clinical aspects of uncomplicated and severe malaria. *Mediterr J Hematol Infect Dis, 4*(1), e2012026. https://doi.org/10.4084/mjhid.2012.026

Bassat, Q., Ordi, J., Vila, J., Ismail, M. R., Carrilho, C., Lacerda, M., . . . Menéndez, C. (2013). Development of a post-mortem procedure to reduce the uncertainty regarding causes of death in developing countries. *The Lancet Global Health, 1*(3), e125-e126. https://doi.org/https://doi.org/10.1016/S2214-109X(13)70037-8

Bayoh, M. N., & Lindsay, S. W. (2003). Effect of temperature on the development of the aquatic stages of Anopheles gambiae sensu stricto (Diptera: Culicidae). *Bulletin of Entomological Research, 93*(5), 375-381. https://doi.org/10.1079/BER2003259

Bayoh, M. N., Walker, E. D., Kosgei, J., Ombok, M., Olang, G. B., Githeko, A. K., . . . Gimnig, J. E. (2014). Persistently high estimates of late night, indoor exposure to malaria vectors

despite high coverage of insecticide treated nets. *Parasites & Vectors, 7*(1), 380. https://doi.org/10.1186/1756-3305-7-380

Beier, J. C. (1998). Malaria Parasite Development in Mosquitoes. *Annual Review of Entomology, 43*(1), 519-543. https://doi.org/10.1146/annurev.ento.43.1.519

Bejon, P., Warimwe, G., Mackintosh, C. L., Mackinnon, M. J., Kinyanjui, S. M., Musyoki, J. N., . . . Marsh, K. (2009). Analysis of Immunity to Febrile Malaria in Children That Distinguishes Immunity from Lack of Exposure. *Infection and Immunity, 77*(5), 1917-1923. https://doi.org/10.1128/IAI.01358-08

Beloconi, A., Nyawanda, B. O., Bigogo, G., Khagayi, S., Obor, D., Danquah, I., . . . Vounatsou, P. (2023). Malaria, climate variability, and interventions: modelling transmission dynamics. *Sci Rep, 13*(1), 7367. https://doi.org/10.1038/s41598-023-33868-8

Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., . . . Gething, P. W. (2015). The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature, advance online publication*. https://doi.org/10.1038/nature15535

Bi, Y., Hu, W., Liu, H., Xiao, Y., Guo, Y., Chen, S., . . . Tong, S. (2012). Can slide positivity rates predict malaria transmission? *Malaria Journal, 11*, 117. https://doi.org/10.1186/1475-2875-11-117

Bigogo, G., Audi, A., Aura, B., Aol, G., Breiman, R. F., & Feikin, D. R. (2010). Health-seeking patterns among participants of population-based morbidity surveillance in rural western Kenya: implications for calculating disease rates. *Int J Infect Dis, 14*(11), e967-973. https://doi.org/10.1016/j.ijid.2010.05.016

Bos, E. (2004). Population and Health in Developing Countries. Volume 1, Population, Health, and Survival at INDEPTH Sites. INDEPTH Network. Ottawa: International Development Research Centre, 2002, pp. 356, (PB) $40.40 ISBN: 0-88936-948-8. *International Journal of Epidemiology, 33*(4), 916-917. https://doi.org/10.1093/ije/dyh220

Bousema, T., Okell, L., Felger, I., & Drakeley, C. (2014). Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nat Rev Microbiol, 12*(12), 833-840. https://doi.org/10.1038/nrmicro3364

Boyce, R. M., Reyes, R., Matte, M., Ntaro, M., Mulogo, E., Lin, F.-C., & Siedner, M. J. (2016). Practical Implications of the Non-Linear Relationship between the Test Positivity Rate and Malaria Incidence. *PLOS ONE, 11*(3), e0152410. https://doi.org/10.1371/journal.pone.0152410

Brooker, S., Akhwale, W. S., Pullan, R. L., Estambale, B. B. A., Clarke, S. E., Snow, R. W., & Hotez, P. J. (2007). Epidemiology of plasmodium-helmith co-infection in Africa: Populations at risk, petential impact on anaemia and prospects for combining control. *The American Journal of Tropical Medicine and Hygiene, 77*(6 Suppl), 88-98.

Butler, S. E., Muok, E. M., Montgomery, S. P., Odhiambo, K., Mwinzi, P. M., Secor, W. E., & Karanja, D. M. (2012, Nov). Mechanism of anemia in Schistosoma mansoni-infected school children in Western Kenya. *American Journal of Tropical Medicine and Hygiene, 87*(5), 862-867. https://doi.org/10.4269/ajtmh.2012.12-0248

Byass, P. (2007). Who needs cause-of-death data? *PLoS Med, 4*(11), e333. https://doi.org/10.1371/journal.pmed.0040333

Byass, P., Chandramohan, D., Clark, S. J., D'Ambruoso, L., Fottrell, E., Graham, W. J., . . . Tollman, S. M. (2012). Strengthening standardised interpretation of verbal autopsy data: the new InterVA-4 tool. *Glob Health Action, 5*, 1-8. https://doi.org/10.3402/gha.v5i0.19281

Byass, P., Herbst, K., Fottrell, E., Ali, M. M., Odhiambo, F., Amek, N., . . . Tollman, S. M. (2015). Comparing verbal autopsy cause of death findings as determined by physician coding and probabilistic modelling: a public health analysis of 54 000 deaths in Africa and Asia. *J Glob Health, 5*(1), 010402. https://doi.org/10.7189/jogh.05.010402

Byass, P., Hussain-Alkhateeb, L., D'Ambruoso, L., Clark, S., Davies, J., Fottrell, E., . . . Petzold, M. (2019). An integrated approach to processing WHO-2016 verbal autopsy data: the InterVA-5 model. *BMC Med, 17*(1), 102. https://doi.org/10.1186/s12916-019-1333-6

Caldrer, S., Ursini, T., Santucci, B., Motta, L., & Angheben, A. (2022, May 13). Soil-Transmitted Helminths and Anaemia: A Neglected Association Outside the Tropics. *Microorganisms, 10*(5). https://doi.org/10.3390/microorganisms10051027

Carneiro, I., Roca-Feltrer, A., Griffin, J. T., Smith, L., Tanner, M., Schellenberg, J. A., . . . Schellenberg, D. (2010). Age-Patterns of Malaria Vary with Severity, Transmission Intensity and Seasonality in Sub-Saharan Africa: A Systematic Review and Pooled Analysis. *PLOS ONE, 5*(2), e8988. https://doi.org/10.1371/journal.pone.0008988

Cavill, I., Auerbach, M., Bailie, G. R., Barrett-Lee, P., Beguin, Y., Kaltwasser, P., . . . Wilson, K. (2006). Iron and the anaemia of chronic disease: a review and strategic recommendations. *Current Medical Research and Opinion, 22*(4), 731-737. https://doi.org/10.1185/030079906X100096

Cibulskis, R. E., Aregawi, M., Williams, R., Otten, M., & Dye, C. (2011). Worldwide Incidence of Malaria in 2009: Estimates, Time Trends, and a Critique of Methods. *PLoS Med, 8*(12), e1001142. https://doi.org/10.1371/journal.pmed.1001142

Clancy, D., & O'Neill, P., D. (2008). Bayesian estimation of the basic reproduction number in stochastic epidemic models. *Bayesian Analysis, 3*(4), 737-757. https://doi.org/10.1214/08-BA328

Corsi, D. J., Neuman, M., Finlay, J. E., & Subramanian, S. V. (2012). Demographic and health surveys: a profile. *International Journal of Epidemiology, 41*(6), 1602-1613. https://doi.org/10.1093/ije/dys184

Cressie, N. A. C. (1993). *Statistics for Spatial Data* (Revised Edition ed.). Wiley-Interscience Publication.

Derman, R. J., & Patted, A. (2023). Overview of iron deficiency and iron deficiency anemia in women and girls of reproductive age. *Int J Gynaecol Obstet, 162 Suppl 2*, 78-82. https://doi.org/10.1002/ijgo.14950

Desai, M., Buff, A. M., Khagayi, S., Byass, P., Amek, N., van Eijk, A., . . . Hamel, M. J. (2014). Age-Specific Malaria Mortality Rates in the KEMRI/CDC Health and Demographic Surveillance System in Western Kenya, 2003–2010. *PLOS ONE, 9*(9), e106197. https://doi.org/10.1371/journal.pone.0106197

Desai, M. R., Terlouw, D. J., Kwena, A. M., Phillips-Howard, P. A., Kariuki, S. K., Wannemuehler, K. A., . . . Ter Kuile, F. O. (2005). Factors associated with hemoglobin concentrations in pre-school children in Western Kenya: cross-sectional studies. *The American Journal of Tropical Medicine and Hygiene, 72*(1), 47-59.

Dhiman, S. (2019). Are malaria elimination efforts on right track? An analysis of gains achieved and challenges ahead. *Infect Dis Poverty, 8*(1), 14. https://doi.org/10.1186/s40249-019-0524-x

Diboulo, E., Sié, A., Diadier, D. A., Voules, D. A. K., Yé, Y., & Vounatsou, P. (2015). Bayesian variable selection in modelling geographical heterogeneity in malaria transmission from sparse data: an application to Nouna Health and Demographic Surveillance System (HDSS) data, Burkina Faso. *Parasites & Vectors, 8*(1), 1-14. https://doi.org/10.1186/s13071-015-0679-7

Diggle, P. J., Tawn, J. A., & Moyeed, R. A. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society Series C-Applied Statistics, 47*, 299-326. https://doi.org/Doi 10.1111/1467-9876.00113

Division of Malaria Control [Ministry of Public Health and Sanitation], Kenya National Bureau of Statistics, & ICF Macro. (2011). *2010 Kenya Malaria Indicator Survey*.

Division of National Malaria Programme (DNMP), & ICF. (2021). *Kenya Malaria Indicator Survey 2020*.

Douglas, N. M., Anstey, N. M., Buffet, P. A., Poespoprodjo, J. R., Yeo, T. W., White, N. J., & Price, R. N. (2012). The anaemia of Plasmodium vivax malaria. *Malaria Journal, 11*(1), 135. https://doi.org/10.1186/1475-2875-11-135

Drakeley, C. J., Corran, P. H., Coleman, P. G., Tongren, J. E., McDonald, S. L. R., Carneiro, I., . . . Riley, E. M. (2005). Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure. *Proceedings of the National Academy of Sciences of the United States of America, 102*(14), 5108-5113. https://doi.org/10.1073/pnas.0408725102

Ecker, M. D., & Gelfand, A. E. (1997). Bayesian Variogram Modeling for an Isotropic Spatial Process. *Journal of Agricultural, Biological, and Environmental Statistics, 2*(4), 347-369. https://doi.org/10.2307/1400508

Ejigu, B. A., Wencheko, E., & Berhane, K. (2018). Spatial pattern and determinants of anaemia in Ethiopia. *PLOS ONE, 13*(5), e0197171. https://doi.org/10.1371/journal.pone.0197171

Feikin, D. R., Adazu, K., Obor, D., Ogwang, S., Vulule, J., Hamel, M. J., & Laserson, K. (2010). Mortality and health among internally displaced persons in western Kenya following post-election violence, 2008: novel use of demographic surveillance. *Bull World Health Organ, 88*(8), 601-608. https://doi.org/10.2471/blt.09.069732

Feikin, D. R., Nguyen, L. M., Adazu, K., Ombok, M., Audi, A., Slutsker, L., & Lindblade, K. A. (2009). The impact of distance of residence from a peripheral health facility on pediatric health utilisation in rural western Kenya. *Trop Med Int Health, 14*(1), 54-61. https://doi.org/10.1111/j.1365-3156.2008.02193.x

Foote, E. M., Sullivan, K. M., Ruth, L. J., Oremo, J., Sadumah, I., Williams, T. N., & Suchdev, P. S. (2013). Determinants of Anemia among Preschool Children in Rural, Western Kenya. *The American Journal of Tropical Medicine and Hygiene, 88*(4), 757-764. https://doi.org/10.4269/ajtmh.12-0560

Fotso, J.-C. (2007). Urban–rural differentials in child malnutrition: Trends and socioeconomic correlates in sub-Saharan Africa. *Health & Place, 13*(1), 205-223. https://doi.org/10.1016/j.healthplace.2006.01.004

Fottrell, E. (2009). Dying to count: mortality surveillance in resource-poor settings. *Glob Health Action, 2*. https://doi.org/10.3402/gha.v2i0.1926

Gargano, J. W., Laserson, K., Muttai, H., Odhiambo, F., Orimba, V., Adamu-Zeh, M., . . . Ackers, M. (2012). The adult population impact of HIV care and antiretroviral therapy in a resource poor setting, 2003-2008. *Aids, 26*(12), 1545-1554. https://doi.org/10.1097/QAD.0b013e328353b7b9

GBD Anaemia Collaborators. (2023). Prevalence, years lived with disability, and trends in anaemia burden by severity and cause, 1990-2021: findings from the Global Burden of Disease Study 2021. *Lancet Haematol*. https://doi.org/10.1016/S2352-3026(23)00160-6

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association, 85*(410), 398-409. https://doi.org/10.2307/2289776

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science, 7*(4), 457-472. http://www.jstor.org/stable/2246093

Gemperli, A., Vounatsou, P., Sogoba, N., & Smith, T. (2006). Malaria Mapping Using Transmission Models: Application to Survey Data from Mali. *American Journal of Epidemiology, 163*(3), 289-297. https://doi.org/10.1093/aje/kwj026

Gething, P. W., Casey, D. C., Weiss, D. J., Bisanzio, D., Bhatt, S., Cameron, E., . . . Lim, S. S. (2016). Mapping Plasmodium falciparum Mortality in Africa between 1990 and 2015.

New England Journal of Medicine, 0(0), null. https://doi.org/10.1056/NEJMoa1606701

Gething, P. W., Noor, A. M., Gikandi, P. W., Ogara, E. A. A., Hay, S. I., Nixon, M. S., . . . Atkinson, P. M. (2006). Improving imperfect data from health management information systems in Africa using space-time geostatistics. *PLOS Medicine, 3*(6), 825-831. https://doi.org/ARTN e271
10.1371/journal.pmed.0030271

Giardina, F., Gosoniu, L., Konate, L., Diouf, M. B., Perry, R., Gaye, O., . . . Vounatsou, P. (2012). Estimating the Burden of Malaria in Senegal: Bayesian Zero-Inflated Binomial Geostatistical Modeling of the MIS 2008 Data. *PLOS ONE, 7*(3), e32625. https://doi.org/10.1371/journal.pone.0032625

Githeko, A. K., Ayisi, J. M., Odada, P. K., Atieli, F. K., Ndenga, B. A., Githure, J. I., & Yan, G. (2006). Topography and malaria transmission heterogeneity in western Kenya highlands: prospects for focal vector control. *Malaria Journal, 5*(1), 107. https://doi.org/10.1186/1475-2875-5-107

Gosoniu, L., Msengwa, A., Lengeler, C., & Vounatsou, P. (2012). Spatially Explicit Burden Estimates of Malaria in Tanzania: Bayesian Geostatistical Modeling of the Malaria Indicator Survey Data. *PLOS ONE, 7*(5), e23966. https://doi.org/10.1371/journal.pone.0023966

Greenland, S., Rothman, K. J., & Lash, T. L. (2008). Measures of effect and measures of association. Chapter 4. In K. J. Rothman, S. Greenland, & T. L. Lash (Eds.), *Modern epidemiology* (4th ed., Vol. 54, pp. 62-67).

Greenwood, B. M., Bojang, K., Whitty, C. J., & Targett, G. A. (2005). Malaria. *The Lancet, 365*(9469), 1487-1498. https://doi.org/10.1016/S0140-6736(05)66420-3

Haider, B. A., Olofin, I., Wang, M., Spiegelman, D., Ezzati, M., & Fawzi, W. W. (2013). Anaemia, prenatal iron use, and risk of adverse pregnancy outcomes: systematic review and meta-analysis. *BMJ, 346*, f3443. https://doi.org/10.1136/bmj.f3443

Hamel, M. J., Adazu, K., Obor, D., Sewe, M., Vulule, J., Williamson, J. M., . . . Laserson, K. F. (2011). A Reversal in Reductions of Child Mortality in Western Kenya, 2003-2009. *The American Journal of Tropical Medicine and Hygiene, 85*(4), 597-605. https://doi.org/10.4269/ajtmh.2011.10-0678

Hanboonkunupakarn, B., & White, N. J. (2022). Advances and roadblocks in the treatment of malaria. *Br J Clin Pharmacol, 88*(2), 374-382. https://doi.org/10.1111/bcp.14474

Hawley, W. A., Phillips-Howard, P. A., ter Kuile, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., . . . Hightower, A. W. (2003). Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *The American Journal of Tropical Medicine and Hygiene, 68*(4 Suppl), 121-127.

Herrera, S., Enuameh, Y., Adjei, G., Ae-Ngibise, K. A., Asante, K. P., Sankoh, O., . . . Yé, Y. (2017). A systematic review and synthesis of the strengths and limitations of measuring

malaria mortality through verbal autopsy. *Malaria Journal, 16*(1), 421. https://doi.org/10.1186/s12936-017-2071-x

Hinne, I. A., Attah, S. K., Mensah, B. A., Forson, A. O., & Afrane, Y. A. (2021). Larval habitat diversity and Anopheles mosquito species distribution in different ecological zones in Ghana. *Parasites & Vectors, 14*(1), 193. https://doi.org/10.1186/s13071-021-04701-w

INDEPTH Network. (2005). *Measuring Health Equity in Small Areas: Findings from Demographic Surveillance Systems* Ashgate Pub Co.

Jensen, T. P., Bukirwa, H., Njama-Meya, D., Francis, D., Kamya, M. R., Rosenthal, P. J., & Dorsey, G. (2009). Use of the slide positivity rate to estimate changes in malaria incidence in a cohort of Ugandan children. *Malaria Journal, 8*, 213. https://doi.org/10.1186/1475-2875-8-213

Kadobera, D., Sartorius, B., Masanja, H., Mathew, A., & Waiswa, P. (2012). The effect of distance to formal health facility on childhood mortality in rural Tanzania, 2005-2007. *Global Health Action, 5*(0). https://doi.org/10.3402/gha.v5i0.19099

Karra, M., Fink, G., & Canning, D. (2016). Facility distance and child mortality: a multi-country study of health facility access, service utilization, and child health outcomes. *International Journal of Epidemiology*, dyw062. https://doi.org/10.1093/ije/dyw062

Kasasa, S., Asoala, V., Gosoniu, L., Anto, F., Adjuik, M., Tindana, C., . . . Vounatsou, P. (2013). Spatio-temporal malaria transmission patterns in Navrongo demographic surveillance site, northern Ghana. *Malaria Journal, 12*(1), 63. https://doi.org/10.1186/1475-2875-12-63

Kassebaum, N. J. (2016). The Global Burden of Anemia. *Hematology/Oncology Clinics of North America, 30*(2), 247-308. https://doi.org/10.1016/j.hoc.2015.11.002

Khagayi, S., Amek, N., Bigogo, G., Odhiambo, F., & Vounatsou, P. (2017). Bayesian spatio-temporal modeling of mortality in relation to malaria incidence in Western Kenya. *PLOS ONE, 12*(7), e0180516. https://doi.org/10.1371/journal.pone.0180516

Khagayi, S., Desai, M., Amek, N., Were, V., Onyango, E. D., Odero, C., . . . Vounatsou, P. (2019). Modelling the relationship between malaria prevalence as a measure of transmission and mortality across age groups. *Malaria Journal, 18*(1), 247. https://doi.org/10.1186/s12936-019-2869-9

Kokaliaris, C., Garba, A., Matuska, M., Bronzan, R. N., Colley, D. G., Dorkenoo, A. M., . . . Vounatsou, P. (2022). Effect of preventive chemotherapy with praziquantel on schistosomiasis among school-aged children in sub-Saharan Africa: a spatiotemporal modelling study. *The Lancet. Infectious Diseases, 22*(1), 136-149. https://doi.org/10.1016/S1473-3099(21)00090-6

Koukounari, A., Estambale, B. B. A., Kiambo Njagi, J., Cundill, B., Ajanga, A., Crudder, C., . . . Brooker, S. (2008). Relationships between anaemia and parasitic infections in Kenyan schoolchildren: A Bayesian hierarchical modelling approach. *International Journal for Parasitology, 38*(14), 1663-1671. https://doi.org/10.1016/j.ijpara.2008.05.013

Koumamba, A. P., Bisvigou, U. J., Ngoungou, E. B., & Diallo, G. (2021). Health information systems in developing countries: case of African countries. *BMC Med Inform Decis Mak, 21*(1), 232. https://doi.org/10.1186/s12911-021-01597-5

Kwambai, T. K., Kariuki, S., Smit, M. R., Nevitt, S., Onyango, E., Oneko, M., . . . Ter Kuile, F. O. (2023). Post-Discharge Risk of Mortality in Children under 5 Years of Age in Western Kenya: A Retrospective Cohort Study. *American Journal of Tropical Medicine and Hygiene*. https://doi.org/10.4269/ajtmh.23-0186

Laishram, D. D., Sutton, P. L., Nanda, N., Sharma, V. L., Sobti, R. C., Carlton, J. M., & Joshi, H. (2012). The complexities of malaria disease manifestations with a focus on asymptomatic malaria. *Malaria Journal, 11*, 29. https://doi.org/10.1186/1475-2875-11-29

Leitao, J., Chandramohan, D., Byass, P., Jakob, R., Bundhamcharoen, K., Choprapawon, C., . . . Mathai, M. (2013). Revising the WHO verbal autopsy instrument to facilitate routine cause-of-death monitoring. *Global Health Action, 6*(0). https://doi.org/10.3402/gha.v6i0.21518

Lo, E., Zhou, G., Oo, W., Afrane, Y., Githeko, A., & Yan, G. (2015). Low Parasitemia in Submicroscopic Infections Significantly Impacts Malaria Diagnostic Sensitivity in the Highlands of Western Kenya. *PLOS ONE, 10*(3), e0121763. https://doi.org/10.1371/journal.pone.0121763

Lubinda, J., Haque, U., Bi, Y., Shad, M. Y., Keellings, D., Hamainza, B., & Moore, A. J. (2021, 2021/06/01/). Climate change and the dynamics of age-related malaria incidence in Southern Africa. *Environmental Research, 197*, 111017. https://doi.org/https://doi.org/10.1016/j.envres.2021.111017

Mafwele, B. J., & Lee, J. W. (2022). Relationships between transmission of malaria in Africa and climate factors. *Scientific Reports, 12*(1), 14392. https://doi.org/10.1038/s41598-022-18782-9

Malhotra, I., Dent, A., Mungai, P., Wamachi, A., Ouma, J. H., Narum, D. L., . . . King, C. L. (2009). Can Prenatal Malaria Exposure Produce an Immune Tolerant Phenotype?: A Prospective Birth Cohort Study in Kenya. *PLoS Med, 6*(7), e1000116. https://doi.org/10.1371/journal.pmed.1000116

Marsh, K., & Snow, R. W. (1999). Malaria transmission and morbidity. *Parassitologia, 41*(1-3), 241-246.

Masaba, B. B., Moturi, J. K., Taiswa, J., & Mmusi-Phetoe, R. M. (2020). Devolution of healthcare system in Kenya: progress and challenges. *Public Health, 189*, 135-140. https://doi.org/10.1016/j.puhe.2020.10.001

Mboera, L. E. G., Senkoro, K. P., Mayala, B. K., Rumisha, S. F., Rwegoshora, R. T., Mlozi, M. R. S., & Shayo, E. H. (2010). Spatio-temporal variation in malaria transmission intensity in five agro-ecosystems in Mvomero district, Tanzania. *Geospatial Health, 4*(2), 167-178.

MEASURE Evaluation, MEASURE DHS, President's Malaria Initiative, Roll Back Malaria Partnership, UNICEF, & World Health Organization. (2013). *Household survey indicators for malaria control*.

Menendez, C., Fleming, A. F., & Alonso, P. L. (2000). Malaria-related anaemia. *Parasitol Today, 16*(11), 469-476. https://doi.org/10.1016/s0169-4758(00)01774-9

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., de Savigny, D., Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet, 386*(10001), 1395-1406. https://doi.org/10.1016/s0140-6736(15)60171-4

Miller, L. H., Baruch, D. I., Marsh, K., & Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature, 415*(6872), 673-679. https://doi.org/10.1038/415673a

Ministry of Health, D. o. N. (2013). *Accelerating reduction of iron deficiency anaemia among pregnant women in Kenya: Plan of Action 2012-2017*. https://www.advancingnutrition.org/resources/anemia-toolkit/kenyas-plan-action-accelerating-reduction-iron-deficiency-anemia-among

Moss, W. J., Dorsey, G., Mueller, I., Laufer, M. K., Krogstad, D. J., Vinetz, J. M., . . . Kazura, J. W. (2015). Malaria Epidemiology and Control within the International Centers of Excellence for Malaria Research. *The American Journal of Tropical Medicine and Hygiene, 93*(3 Suppl), 5-15. https://doi.org/10.4269/ajtmh.15-0006

Müller, O., & Krawinkel, M. (2005). Malnutrition and health in developing countries. *CMAJ : Canadian Medical Association Journal, 173*(3), 279-286. https://doi.org/10.1503/cmaj.050342

Murray, C. J., Ortblad, K. F., Guinovart, C., Lim, S. S., Wolock, T. M., Roberts, D. A., . . . Vos, T. (2014). Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet, 384*(9947), 1005-1070. https://doi.org/10.1016/S0140-6736(14)60844-8

Murray, C. J., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., . . . Lopez, A. D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *The Lancet, 379*(9814), 413-431. https://doi.org/10.1016/S0140-6736(12)60034-8

Mwangi, T. W., Bethony, J. M., & Brooker, S. (2006). Malaria and helminth interactions in humans: an epidemiological viewpoint. *Ann Trop Med Parasitol, 100*(7), 551-570. https://doi.org/10.1179/136485906x118468

Nacher, M. (2011). Interactions between worms and malaria: good worms or bad worms? *Malar J, 10*, 259. https://doi.org/10.1186/1475-2875-10-259

NASA JPL. *NASA Land Data Products and Services*. https://lpdaac.usgs.gov/product_search/

National AIDS and STI Control Programme (NASCOP). (2020). *Preliminary KENPHIA 2018 Report.*

National AIDS and STI Control Programme (NASCOP), K. (2014). *Kenya AIDS Indicator Survey 2012: Final Report*.

Ngesa, O., & Mwambi, H. (2014). Prevalence and Risk Factors of Anaemia among Children Aged between 6 Months and 14 Years in Kenya. *PLOS ONE, 9*(11), e113756. https://doi.org/10.1371/journal.pone.0113756

Nichols, E. K., Byass, P., Chandramohan, D., Clark, S. J., Flaxman, A. D., Jakob, R., . . . on behalf of the, W. H. O. V. A. W. G. (2018). The WHO 2016 verbal autopsy instrument: An international standard suitable for automated analysis by InterVA, InSilicoVA, and Tariff 2.0. *PLOS Medicine, 15*(1), e1002486. https://doi.org/10.1371/journal.pmed.1002486

Noor, A. M., Gething, P. W., Alegana, V. A., Patil, A. P., Hay, S. I., Muchiri, E., . . . Snow, R. W. (2009). The risks of malaria infection in Kenya in 2009. *BMC Infectious Diseases, 9*(1), 180. https://doi.org/10.1186/1471-2334-9-180

Noor, A. M., Kirui, V. C., Brooker, S. J., & Snow, R. W. (2009). The use of insecticide treated nets by age: implications for universal coverage in Africa. *BMC Public Health, 9*, 369. https://doi.org/10.1186/1471-2458-9-369

Nyawanda, B. O., Beloconi, A., Khagayi, S., Bigogo, G., Obor, D., Otieno, N. A., . . . Vounatsou, P. (2023). The relative effect of climate variability on malaria incidence after scale-up of interventions in western Kenya: A time-series analysis of monthly incidence data from 2008 to 2019. *Parasite Epidemiol Control, 21*, e00297. https://doi.org/10.1016/j.parepi.2023.e00297

O'Hara, R. B., & Sillanpää, M. J. (2009). A review of Bayesian variable selection methods: what, how and which. *Bayesian Analysis, 4*(1), 85-117. https://doi.org/10.1214/09-BA403

O'Meara, W. P., Bejon, P., Mwangi, T. W., Okiro, E. A., Peshu, N., Snow, R. W., . . . Marsh, K. (2008). Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. *The Lancet, 372*(9649), 1555-1562. https://doi.org/10.1016/S0140-6736(08)61655-4

Obonyo, C. O., Vulule, J., Akhwale, W. S., & Grobbee, D. E. (2007). In-Hospital Morbidity and Mortality Due to Severe Malarial Anemia in Western Kenya. *The American Journal of Tropical Medicine and Hygiene, 77*(6 Suppl), 23-28.

Obsomer, V., Defourny, P., & Coosemans, M. (2007). The Anopheles dirus complex: spatial distribution and environmental drivers. *Malar J, 6*, 26. https://doi.org/10.1186/1475-2875-6-26

Odhiambo, F. O., Laserson, K. F., Sewe, M., Hamel, M. J., Feikin, D. R., Adazu, K., . . . Vulule, J. M. (2012). Profile: The KEMRI/CDC Health and Demographic Surveillance System-

-Western Kenya. *International Journal of Epidemiology, 41*(4), 977-987. https://doi.org/10.1093/ije/dys108

Odiere, M. R., Opisa, S., Odhiambo, G., Jura, W. G., Ayisi, J. M., Karanja, D. M., & Mwinzi, P. N. (2011). Geographical distribution of schistosomiasis and soil-transmitted helminths among school children in informal settlements in Kisumu City, Western Kenya. *Parasitology, 138*(12), 1569-1577. https://doi.org/10.1017/s003118201100059x

Oduma, C. O., Ombok, M., Zhao, X., Huwe, T., Ondigo, B. N., Kazura, J. W., . . . Koepfli, C. (2023). Altitude, not potential larval habitat availability, explains pronounced variation in Plasmodium falciparum infection prevalence in the western Kenya highlands. *PLOS Glob Public Health, 3*(4), e0001505. https://doi.org/10.1371/journal.pgph.0001505

Okello, G., Molyneux, S., Zakayo, S., Gerrets, R., & Jones, C. (2019). Producing routine malaria data: an exploration of the micro-practices and processes shaping routine malaria data quality in frontline health facilities in Kenya. *Malaria Journal, 18*(1), 420. https://doi.org/10.1186/s12936-019-3061-y

Okiro, E. A., Alegana, V. A., Noor, A. M., Mutheu, J. J., Juma, E., & Snow, R. W. (2009). Malaria paediatric hospitalization between 1999 and 2008 across Kenya. *BMC Medicine, 7*, 75. https://doi.org/10.1186/1741-7015-7-75

Ombok, M., Adazu, K., Odhiambo, F., Bayoh, N., Kiriinya, R., Slutsker, L., . . . Feikin, D. R. (2010). Geospatial distribution and determinants of child mortality in rural western Kenya 2002-2005. *Trop Med Int Health, 15*(4), 423-433. https://doi.org/10.1111/j.1365-3156.2010.02467.x

Ondeto, B. M., Wang, X., Atieli, H., Zhong, D., Zhou, G., Lee, M.-C., . . . Yan, G. (2022). A prospective cohort study of Plasmodium falciparum malaria in three sites of Western Kenya. *Parasites & Vectors, 15*(1), 416. https://doi.org/10.1186/s13071-022-05503-4

Otambo, W. O., Omondi, C. J., Ochwedo, K. O., Onyango, P. O., Atieli, H., Lee, M. C., . . . Kazura, J. (2022). Risk associations of submicroscopic malaria infection in lakeshore, plateau and highland areas of Kisumu County in western Kenya. *PLOS ONE, 17*(5), e0268463. https://doi.org/10.1371/journal.pone.0268463

Ouma, P., Eijk, A. M. V., Hamel, M. J., Parise, M., Ayisi, J. G., Otieno, K., . . . 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. https://doi.org/10.1111/j.1365-3156.2007.01960.x

Phillips, M. A., Burrows, J. N., Manyando, C., van Huijsduijnen, R. H., Van Voorhis, W. C., & Wells, T. N. C. (2017). Malaria. *Nat Rev Dis Primers, 3*, 17050. https://doi.org/10.1038/nrdp.2017.50

Rakislova, N., Fernandes, F., Lovane, L., Jamisse, L., Castillo, P., Sanz, A., . . . Ordi, J. (2019). Standardization of Minimally Invasive Tissue Sampling Specimen Collection and Pathology Training for the Child Health and Mortality Prevention Surveillance

Network. *Clinical Infectious Diseases, 69*(Supplement_4), S302-S310. https://doi.org/10.1093/cid/ciz565

Rakislova, N., Jordao, D., Ismail, M. R., Mayor, A., Cisteró, P., Marimon, L., . . . Ordi, J. (2021). Accuracy of verbal autopsy, clinical data and minimally invasive autopsy in the evaluation of malaria-specific mortality: an observational study. *BMJ Glob Health, 6*(6). https://doi.org/10.1136/bmjgh-2021-005218

Rek, J., Blanken, S. L., Okoth, J., Ayo, D., Onyige, I., Musasizi, E., . . . Conrad, M. D. (2022). Asymptomatic School-Aged Children Are Important Drivers of Malaria Transmission in a High Endemicity Setting in Uganda. *The Journal of Infectious Diseases, 226*(4), 708-713. https://doi.org/10.1093/infdis/jiac169

Robert, C., & Casella, G. (2011). A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data. *Statistical Science, 26*(1), 102-115. https://doi.org/10.1214/10-STS351

Roll Back Malaria Monitoring & Evaluation Reference Group. (2018). *Household survey indicators for malaria control*. https://www.malariasurveys.org/documents/Household%20Survey%20Indicators%20for%20Malaria%20Control_FINAL.pdf

Roser, M., & Ritchie, H. (2019). *"Malaria". Published online at OurWorldInData.org*. https://ourworldindata.org/malaria

Ross, A., & Smith, T. (2006). The Effect of Malaria Transmission Intensity on Neonatal Mortality in Endemic Areas. *The American Journal of Tropical Medicine and Hygiene, 75*(2 suppl), 74-81.

Rotejanaprasert, C., Lee, D., Ekapirat, N., Sudathip, P., & Maude, R. J. (2021). Spatiotemporal distributed lag modelling of multiple Plasmodium species in a malaria elimination setting. *Statistical Methods in Medical Research, 30*(1), 22-34. https://doi.org/10.1177/0962280220938977

Rumisha, S. F., Smith, T. A., Masanja, H., Abdulla, S., & Vounatsou, P. (2014). Relationship between child survival and malaria transmission: an analysis of the malaria transmission intensity and mortality burden across Africa (MTIMBA) project data in Rufiji demographic surveillance system, Tanzania. *Malaria Journal, 13*(1), 124. https://doi.org/10.1186/1475-2875-13-124

Rutherford, M. (2009). Access to health care and mortality of children under 5 years of age in the Gambia: a case-control study. *Bulletin of the World Health Organization, 87*(3), 216-225. https://doi.org/10.2471/BLT.08.052175

Ryan, S. J., Lippi, C. A., & Zermoglio, F. (2020). Shifting transmission risk for malaria in Africa with climate change: a framework for planning and intervention. *Malaria Journal, 19*(1), 170. https://doi.org/10.1186/s12936-020-03224-6

Sachs, J., & Malaney, P. (2002). The economic and social burden of malaria. *Nature, 415*(6872), 680-685. https://doi.org/10.1038/415680a

Safiri, S., Kolahi, A.-A., Noori, M., Nejadghaderi, S. A., Karamzad, N., Bragazzi, N. L., . . . Grieger, J. A. (2021). Burden of anemia and its underlying causes in 204 countries and territories, 1990–2019: results from the Global Burden of Disease Study 2019. *Journal of Hematology & Oncology, 14*(1), 185. https://doi.org/10.1186/s13045-021-01202-2

Sankaran, V. G., & Weiss, M. J. (2015). Anemia: progress in molecular mechanisms and therapies. *Nature Medicine, 21*(3), 221-230. https://doi.org/10.1038/nm.3814

Sankoh, O., & Binka, F. (2005). INDEPTH Network: a viable platform for the assessment of malaria risk in developing countries. In *Environmental Change and Malaria Risk: Global and Local Implications* (Vol. 9, pp. 99-105). Frontis.

Sankoh, O., & Byass, P. (2012). The INDEPTH Network: filling vital gaps in global epidemiology. *International Journal of Epidemiology, 41*(3), 579-588. https://doi.org/10.1093/ije/dys081

Schoeps, A., Gabrysch, S., Niamba, L., Sié, A., & Becher, H. (2011). The effect of distance to health-care facilities on childhood mortality in rural Burkina Faso. *Am J Epidemiol, 173*(5), 492-498. https://doi.org/10.1093/aje/kwq386

Sedda, L., Qi, Q., & Tatem, A. J. (2015). A geostatistical analysis of the association between armed conflicts and Plasmodium falciparum malaria in Africa, 1997–2010. *Malaria Journal, 14*, 500. https://doi.org/10.1186/s12936-015-1024-5

Setel, P. W., Macfarlane, S. B., Szreter, S., Mikkelsen, L., Jha, P., Stout, S., & AbouZahr, C. (2007). A scandal of invisibility: making everyone count by counting everyone. *Lancet, 370*(9598), 1569-1577. https://doi.org/10.1016/s0140-6736(07)61307-5

Sewe, M. O., Rocklöv, J., Williamson, J., Hamel, M., Nyaguara, A., Odhiambo, F., & Laserson, K. (2015). The Association of Weather Variability and Under Five Malaria Mortality in KEMRI/CDC HDSS in Western Kenya 2003 to 2008: A Time Series Analysis. *International Journal of Environmental Research and Public Health, 12*(2), 1983-1997. https://doi.org/10.3390/ijerph120201983

Smith, T., Leuenberger, R., & Lengeler, C. (2001). Child mortality and malaria transmission intensity in Africa. *Trends in parasitology, 17*(3), 145-149. https://doi.org/10.1016/S1471-4922(00)01814-6

Snow, R. W. (2014). Sixty years trying to define the malaria burden in Africa: have we made any progress? *BMC Medicine, 12*(1), 227. https://doi.org/10.1186/s12916-014-0227-x

Snow, R. W., & Omumbo, J. A. (2006). Malaria. In D. T. Jamison, R. G. Feachem, M. W. Makgoba, E. R. Bos, F. K. Baingana, K. J. Hofman, & K. O. Rogo (Eds.), *Disease and Mortality in Sub-Saharan Africa* (2nd ed.). World Bank.

Soares Magalhães, R. J., & Clements, A. C. A. (2011). Mapping the Risk of Anaemia in Preschool-Age Children: The Contribution of Malnutrition, Malaria, and Helminth Infections in West Africa. *PLoS Med, 8*(6), e1000438. https://doi.org/10.1371/journal.pmed.1000438

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583-639. https://doi.org/10.1111/1467-9868.00353

Ssempiira, J., Kissa, J., Nambuusi, B., Mukooyo, E., Opigo, J., Makumbi, F., . . . Vounatsou, P. (2018). Interactions between climatic changes and intervention effects on malaria spatio-temporal dynamics in Uganda. *Parasite Epidemiology and Control, 3*(3), e00070. https://doi.org/10.1016/j.parepi.2018.e00070

Ssempiira, J., Nambuusi, B., Kissa, J., Agaba, B., Makumbi, F., Kasasa, S., & Vounatsou, P. (2017). The contribution of malaria control interventions on spatio-temporal changes of parasitaemia risk in Uganda during 2009-2014. *Parasit Vectors, 10*(1), 450. https://doi.org/10.1186/s13071-017-2393-0

Stanisic, D. I., Fowkes, F. J. I., Koinari, M., Javati, S., Lin, E., Kiniboro, B., . . . Beeson, J. G. (2014). Acquisition of antibodies against Plasmodium falciparum merozoites and malaria immunity in young children: influence of age, force of infection, and magnitude of response. *Infection and Immunity*, IAI.02398-02314. https://doi.org/10.1128/IAI.02398-14

Stevens, G. A., Finucane, M. M., De-Regil, L. M., Paciorek, C. J., Flaxman, S. R., Branca, F., . . . Ezzati, M. (2013). Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995–2011: a systematic analysis of population-representative data. *The Lancet. Global Health, 1*(1), e16-e25. https://doi.org/10.1016/S2214-109X(13)70001-9

Streatfield, P. K., Khan, W. A., Bhuiya, A., Hanifi, S. M., Alam, N., Diboulo, E., . . . Byass, P. (2014). Malaria mortality in Africa and Asia: evidence from INDEPTH health and demographic surveillance system sites. *Glob Health Action, 7*, 25369. https://doi.org/10.3402/gha.v7.25369

Sumbele, I. U. N., Samje, M., & Nkuo-Akenji, T. (2013). A longitudinal study on anaemia in children with Plasmodium falciparum infection in the Mount Cameroon region: prevalence, risk factors and perceptions by caregivers. *BMC Infectious Diseases, 13*(1), 123. https://doi.org/10.1186/1471-2334-13-123

Terekeci, H. M., Kucukardali, Y., Onem, Y., Erikci, A. A., Kucukardali, B., Sahan, B., . . . Oktenli, C. (2010). Relationship between anaemia and cognitive functions in elderly people. *European Journal of Internal Medicine, 21*(2), 87-90. https://doi.org/10.1016/j.ejim.2009.12.005

Thomas, L. M., D'Ambruoso, L., & Balabanova, D. (2018). Verbal autopsy in health policy and systems: a literature review. *BMJ Glob Health, 3*(2), e000639. https://doi.org/10.1136/bmjgh-2017-000639

Villena, O. C., Ryan, S. J., Murdock, C. C., & Johnson, L. R. (2022). Temperature impacts the environmental suitability for malaria transmission by Anopheles gambiae and Anopheles stephensi. *Ecology, 103*(8), e3685. https://doi.org/10.1002/ecy.3685

Walldorf, J. A., Cohee, L. M., Coalson, J. E., Bauleni, A., Nkanaunena, K., Kapito-Tembo, A., . . . Laufer, M. K. (2015). School-Age Children Are a Reservoir of Malaria Infection in Malawi. *PLOS ONE, 10*(7). https://doi.org/10.1371/journal.pone.0134061

Wang, H., He, L., Yin, J., Yu, Z., Liu, S., & Yan, D. (2022). Effects of Effective Precipitation and Accumulated Temperature on the Terrestrial EVI (Enhanced Vegetation Index) in the Yellow River Basin, China. *Atmosphere, 13*(10), 1555. https://www.mdpi.com/2073-4433/13/10/1555

Weatherall, D. J., Miller, L. H., Baruch, D. I., Marsh, K., Doumbo, O. K., Casals-Pascual, C., & Roberts, D. J. (2002). Malaria and the red cell. *Hematology Am Soc Hematol Educ Program*, 35-57. https://doi.org/10.1182/asheducation-2002.1.35

Weiss, D. J., Bertozzi-Villa, A., Rumisha, S. F., Amratia, P., Arambepola, R., Battle, K. E., . . . Gething, P. W. (2021). Indirect effects of the COVID-19 pandemic on malaria intervention coverage, morbidity, and mortality in Africa: a geospatial modelling analysis. *Lancet Infect Dis, 21*(1), 59-69. https://doi.org/10.1016/s1473-3099(20)30700-3

Weiss, D. J., Lucas, T. C. D., Nguyen, M., Nandi, A. K., Bisanzio, D., Battle, K. E., . . . Gething, P. W. (2019). Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study. *Lancet, 394*(10195), 322-331. https://doi.org/10.1016/s0140-6736(19)31097-9

Were, V., Buff, A. M., Desai, M., Kariuki, S., Samuels, A., ter Kuile, F. O., . . . Niessen, L. (2018). Socioeconomic health inequality in malaria indicators in rural western Kenya: evidence from a household malaria survey on burden and care-seeking behaviour. *Malaria Journal, 17*(1), 166. https://doi.org/10.1186/s12936-018-2319-0

White, N. J. (2018). Anaemia and malaria. *Malaria Journal, 17*(1), 371. https://doi.org/10.1186/s12936-018-2509-9

White, N. J., Pukrittayakamee, S., Hien, T. T., Faiz, M. A., Mokuolu, O. A., & Dondorp, A. M. (2014). Malaria. *The Lancet, 383*(9918), 723-735. https://doi.org/https://doi.org/10.1016/S0140-6736(13)60024-0

Whitworth, J., Morgan, D., Quigley, M., Smith, A., Mayanja, B., Eotu, H., . . . Ojwiya, A. (2000). Effect of HIV-1 and increasing immunosuppression on malaria parasitaemia and clinical episodes in adults in rural Uganda: a cohort study. *The Lancet, 356*(9235), 1051-1056. https://doi.org/10.1016/S0140-6736(00)02727-6

WHO Malaria Policy Advisory Committee and Secretariat. (2012). Inaugural meeting of the malaria policy advisory committee to the WHO: conclusions and recommendations. *Malar J, 11*, 137. https://doi.org/10.1186/1475-2875-11-137

WHO Malaria Policy Advisory Committee and Secretariat. (2013). *Malaria Policy Advisory Committee to the WHO: conclusions and recommendations of the March 2013 meeting*.

Wiegand, R. E., Mwinzi, P. N. M., Montgomery, S. P., Chan, Y. L., Andiego, K., Omedo, M., . . . Secor, W. E. (2017). A Persistent Hotspot of Schistosoma mansoni Infection in a Five-Year Randomized Trial of Praziquantel Preventative Chemotherapy Strategies. *J Infect Dis, 216*(11), 1425-1433. https://doi.org/10.1093/infdis/jix496

World Bank. *Prevalence of anemia among non-pregnant women (% of women ages 15-49) - Kenya*. Retrieved 17/08/2023 from https://data.worldbank.org/indicator/SH.ANM.NPRG.ZS?locations=KE

World Health, O. (2000). Severe falciparum malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 94*, 1-90. https://doi.org/https://doi.org/10.1016/S0035-9203(00)90300-6

World Health Organization. (2011). *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System*.

World Health Organization. (2015a). *The global prevalence of anaemia in 2011*.

World Health Organization. (2015b). *World Malaria Report 2015*.

World Health Organization. (2016). *World Malaria Report 2016*.

World Health Organization. (2017a). *The Double Burden of Malnutrition. Policy brief.* https://www.who.int/publications/i/item/WHO-NMH-NHD-17.3

World Health Organization. (2017b). *Nutritional anaemias: tools for effective prevention and control*.

World Health Organization. (2021). *Global Technical Strategy for Malaria 2016–2030, 2021 update*. https://www.who.int/publications/i/item/9789240031357

World Health Organization. (2022). *World malaria report 2022*.

Yang, F., Liu, X., & Zha, P. (2018). Trends in Socioeconomic Inequalities and Prevalence of Anemia Among Children and Nonpregnant Women in Low- and Middle-Income Countries. *JAMA Netw Open, 1*(5), e182899. https://doi.org/10.1001/jamanetworkopen.2018.2899

Ye, Y., Wamukoya, M., Ezeh, A., Emina, J. B. O., & Sankoh, O. (2012). Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Sahara Africa? *BMC Public Health, 12*(1), 741. https://doi.org/10.1186/1471-2458-12-741

Zhou, G., Afrane, Y. A., Vardo-Zalik, A. M., Atieli, H., Zhong, D., Wamae, P., . . . Yan, G. (2011). Changing Patterns of Malaria Epidemiology between 2002 and 2010 in Western Kenya: The Fall and Rise of Malaria. *PLOS ONE, 6*(5), e20318. https://doi.org/10.1371/journal.pone.0020318

Zhou, G., Li, J. S., Ototo, E. N., Atieli, H. E., Githeko, A. K., & Yan, G. (2014). Evaluation of universal coverage of insecticide-treated nets in western Kenya: field surveys. *Malaria Journal, 13*(1), 351. https://doi.org/10.1186/1475-2875-13-351

Zhou, G., Minakawa, N., Githeko, A. K., & Yan, G. (2004). Association between climate variability and malaria epidemics in the East African highlands. *Proceedings of the National Academy of Sciences of the United States of America, 101*(8), 2375-2380. https://doi.org/10.1073/pnas.0308714100

Zhou, Z., Mitchell, R. M., Kariuki, S., Odero, C., Otieno, P., Otieno, K., . . . Shi, Y. P. (2016). Assessment of submicroscopic infections and gametocyte carriage of Plasmodium falciparum during peak malaria transmission season in a community-based cross-sectional survey in western Kenya, 2012. *Malaria Journal, 15*, 421. https://doi.org/10.1186/s12936-016-1482-4

Zimmerman, P. A., & Howes, R. E. (2015). Malaria diagnosis for malaria elimination. *Current Opinion in Infectious Diseases, 28*(5). https://journals.lww.com/co-infectiousdiseases/fulltext/2015/10000/malaria_diagnosis_for_malaria_elimination.7.aspx