Assessing malaria attributed mortality in west and southern Africa

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät der
Universität Basel

von
Simon Kasasa
aus Kampala, Uganda

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von
Prof. Dr. M. Tanner, PD Dr. P. Vounatsou, Prof. Dr. T. Smith und Prof. John J. Aponte

Basel, den 16, September 2012

Prof. Dr. Jörg Schibler

Dekan

Dedicated to my parents, Mr. & Mrs. Jerome Ddamba

and

my grandmother Ms. Susan Nakiganda

Table of Contents

Acronyms

ADDS	African Data Dissemination Service
AIC	Akaike's information criterion
CDC	Centers for disease control and prevention
DDT	Dichlorodiphenyltrichloroethane
DHS	Demographic and health surveys
EIR	Entomological inoculation rate
EVI	Enhanced Vegetation Index
GPS	Global Positioning Systems
GZIB	Geostatistical zero-inflated binomial
HDSS	Health and demographic surveillance systems
HLC	Human landing catches
HR	Hazard rate
ICD	International Classification of Diseases
IPT	Intermittent preventive treatment
IRS	Indoor residual spraying
ITN	Insect treated nets
LST	Land surface temperature
MARA	Mapping Malaria Risk in Africa
MCMC	Markov chain Monte Carlo
MDG	Millennium Development Goals
MODIS	Moderate Resolution Imaging Spectroradiometer
MR	Mortality rate
MTIMBA	Malaria Transmission Intensity and mortality Burden across Africa
NDVI	Normalized Difference Vegetation Index
NHDSS	Navrongo Health and Demographic surveillance System
PCA	Principal components analysis
RBM	Roll Back Malaria
RF	Rainfall estimates
SES	Socio-Economic Status
SR	Sporozoite rate
SSA	Sub-Saharan Africa
VA	Verbal Autopsy
WHO	World Health Organization
ZIB	Zero-inflated binomial
ZINB	Zero-inflated negative binomial

List of figures

List of tables

Summary

Malaria has persistently remained a serious health and socio-economic problem in developing nations particularly in Sub-Saharan Africa (SSA). There are approximately 500 million cases of malaria each year and close to one million deaths occurring mainly among children under five years. Developing countries spend a reasonable proportion of their gross domestic product (GDP) on malaria which in the end hinders their levels of development.

World Health Organizations (WHO) and partners through the Roll Back Malaria initiative (RBM) have targeted vector control, health promotion and case management (using rapid diagnostic tests and treatment with Artemisinin combination therapy) in order reduce malaria morbidity and mortality cases. Since 2002, funds for promoting malaria control activities have increased exponentially in SSA. Major donors include presidential malaria initiative (PMI) and Global fund to fight AIDS, tuberculosis and malaria (GFATM). Countries which have scaled up the recommended malaria control strategies such as insecticides-treat net (ITN) and treatment of confirmed cases have reported a decline in both morbidity and mortality especially among children. However, these statistics are based on health facilities data and yet in most developing countries many deaths occur at home and are never recorded due to inefficient vital registration systems. Monitoring the progress of such interventions requires reliable sources of data on both the transmission and infection outcome.

In malaria endemic areas, people acquire natural immunity during the early years of their life after getting exposed to repeated infections. This is observed from the reductions in the number of severe malaria-related morbidity and mortality cases especially in children >5 years. Due to the current undertakings that are aimed at reducing malaria exposure, there are concerns about

shifting the disease burden to older children but the required to data to monitor this are not readily available in SSA. Low income countries have resorted to health and demographic surveillance systems (HDSS) to monitor routinely population changes and health outcomes within a defined geographical area.

In 2000, the INDEPTH, a network of HDSS integrated the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project into selected sites' routine activities in order to assess the transmission-malaria mortality relationship taking into account the current interventions. Mortality data and other demographic characteristics were extracted from routinely collected HDSS databases. The entomological data were collected every fortnight from randomly sampled compounds over the 3 years MTIMBA period.

The MTIMBA project generated large geostatistical data that are correlated in space and time. Furthermore, the project captured longitudinal mosquito data that were characterized by many zeros especially during the dry periods. The zeros are due empty traps from a compound or when all the captured mosquitoes are not infectious. Appropriate data analysis therefore should apply models that account for spatial-temporal correlation and the excess zeros in order to avoid over or underestimation of parameters. Zero-inflated geostatistical models account for spatial-temporal correlation by introducing location-specific and time interval random effects which creates more parameters to estimate. Bayesian models implemented via Markov chain Monte Carlo simulation (MCMC) addresses fit of highly parameterized models.

This work applied zero-inflated Bayesian models to estimate malaria attributable mortality across all age-groups using large, correlated and sparse data collected from Navrongo and Manhiça HDSS between 2001 and 2004. The contributions of this thesis were (i) the description of the HDSS data characteristics and relevant methods for analysis; (ii) the spatially

explicit estimates of malaria transmission intensity at monthly intervals; and (iii) the relationship between all-cause mortality and malaria transmission intensity across all age categories.

Chapter 2 described the characteristics of the MTIMBA data. These are large geostatistical, temporal, seasonal and zero-inflated data. The mortality and mosquito data were misaligned because they were captured at different compounds and time periods. Zero-inflated Bayesian spatio-temporal models are the state-of-art in handling such data. The rigorous statistical process was demonstrated by modelling sporozoite rate (SR) data from Manhiça HDSS. The analysis of the MTIMBA data was used as an avenue for building SSA capacity through course work, seminars and mentorship. Site-specific analyses are still on-going. However, the project generated data that is relevant for assessing within and between site malaria transmission heterogeneity.

The Navrongo malaria exposure surfaces described in chapter 3 were obtained from zero-inflated geostatistical models fitting separately the binomial SR data and negative binomial count data by mosquito species. All the models included space and time correlation in addition to the Climate, environmental and seasonality covariates. The entomological inoculation rate (EIR) estimates were derived as a product of predicted man biting rate and SR. Observed EIR in this district was >100 infective bites/person/year. Distance to water to bodies, day temperatures and vegetation were the main predictors of mosquito densities for the two species. The EIR maps clearly indicated that the temporal heterogeneity was stronger than the spatial variation in this area. The same situation was also observed from the analyses of the two MTIMBA sites of Rufiji (Tanzania) and Kisumu (Kenya).

Monthly malaria exposure surfaces (chapter 3) were linked to the nearest compounds where mortality was observed as described in chapter 4. Time to death data were split at monthly

intervals in order to generate Bernoulli and binomial data that were modelled via logistic regression formulations. Spatio-temporal models were fitted to obtain age-specific mortality risk estimates. The model considered 2 covariates; natural logarithm transformed EIR estimates with their measurement errors and age. ITN variable was only included in neonates, post-neonates and child models. The analysis showed a positive log-linear relationship between all-cause mortality and malaria exposure in all the age groups but the association was only important among children (1-4 years) and people >= 60 years. ITN use showed a protective effect among all the under five children, confirming what was observed in Rufiji and Kisumu HDSS.

The methods used in estimating malaria exposure surfaces and mortality risks in chapters 3 and 4 were extended to Manhiça HDSS (Mozambique) data to describe the mortality-malaria transmission relationship for this area (chapter 5). The spatio-temporal age-specific models considered EIR estimates with their measurement errors (to account for the predictive uncertainty) and age as model covariates.

The distance to the nearest water bodies was the only important common predictor of A*n. funestus* and A*n. gambiae* mosquito densities. Malaria transmission intensity declined consistently in this area. The Model-based results indicated a positive log-linear relationship between all-cause mortality and malaria exposure across all age groups namely; the neonates (0-28 days), post-neonates (1-11months), children (1-4years), young people (5-14 years), adults (15- 59years) and old age (>=60 years).

This work contributes to further understand of malaria-mortality relationships. A positive association between mortality and malaria exposure among the under fives is consistent with what was reported from the MTIMBA sites of Rufiji and Kisumu. Completion of the remaining site-specific analyses followed by a meta-analysis will make a great contribution to malaria

epidemiology. Further work however, should consider cohort analysis in order to ascertain whether malaria control interventions have caused a shift in the age of acquired immunity.

Zusammenfassung

Malaria ist nach wie vor ein ernstzunehmendes gesundheitliches und sozioökonomisches Problem in Entwicklungsländern, insbesondere in Subsahara-Afrika (SSA). Jedes Jahr werden ca. 500 Millionen Malariafälle und rund eine Million Todesfälle, hauptsächlich Kinder unter fünf Jahre, gezählt. Ein Großteil des Bruttoinlandsprodukts in Entwicklungsländern fließt in die Bekämpfung von Malaria und kann somit nicht in andere Bereiche zur Entwicklung investiert werden.

Die Weltgesundheitsorganisation (WHO) und die Roll Back Malaria Partnerschaft (RBM) haben sich Vektorkontrolle, Gesundheitsförderung und Fallmanagement (unter Verwendung von schnellen Diagnosetests mit Artemisinin-basierter Kombinationstherapie) zum Ziel gesetzt, um Malariamorbidität als auch -mortaliät zu reduzieren. Seit 2002 sind die Geldmittel für Malariakontrolle in SSA exponentiell gestiegen. Zu den Hauptinvestoren zählen die President's Malaria Initiative (PMI) und Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM). Länder, welche die empfohlenen Strategien zur Malariakontrolle, wie zum Beispiel Insektizid-behandelte Moskitonetze und Behandlung von bestätigten Fällen, erhöht haben, meldeten einen Rückgang sowohl in Morbidität als auch in Mortalität insbesondere unter Kindern. Jedoch basieren diese Statistiken auf Daten von Gesundheitseinrichtungen, wobei in den meisten Entwicklungsländern viele Todesfälle zuhause eintreten und somit aufgrund von ineffizienten Meldewesen nicht registriert werden. Monitorings der Fortschritts solcher Interventionen erfordert zuverlässige Datenquellen bzgl. der Übertragung als auch die Entwicklung der Infektion.

In Malaria-endemischen Gebieten wird die natürliche Immunität nach wiederholten Infektionen in den frühen Lebensjahren erworben. Dies wurde aus der zurückgehenden Zahl der malariabedingten Morbidität und Mortalität, insbesondere bei Kindern unter fünf Jahren, geschlossen. Aufgrund der momentan Initiativen zur Bekämpfung von Malaria herrscht Besorgnis darüber, dass das Risiko der Krankheit auf ältere Kinder überlagert werden könnte. Jedoch gibt es dafür aufgrund mangelnder Daten in SSA bisher keine Belege. Einkommensschwache Länder haben auf Gesundheits- und demographische Überwachungssysteme zurückgegriffen (HDSS) um regelmäßig Veränderungen der Bevölkerung und die gesundheitliche Situation in ausgewählten geographischen Gebieten zu kontrollieren.

In 2000 hat INDEPTH, ein HDSS Netzwerk, das Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) Projekt ins Leben gerufen, um die Beziehung zwischen Malariaübertragung und Mortalität unter Berücksichtigung der momentanen Interventionen zu beurteilen. Mortalitätsdaten und weitere demographische Kennzahlen wurden von der regelmäßig angepassten HDSS Datenbank extrahiert. Entomologische Daten wurden drei Jahre lang in zweiwöchigen Abständen von zufällig ausgewählten Gebieten gesammelt.

Dank des MTIMBA Projekts wurden große geostatistische Daten generiert, welche in Raum und Zeit korreliert sind. Des Weiteren umfasste das Projekt longitudinale Daten bzgl. Moskitos, welche durch zahlreiche Nullwerte, insbesondere während der Trockenperioden, charakterisiert sind. Die Nullwerte entstehen durch Vorliegen von leeren Moskitofallen in einem Gebiet oder wenn keine der gefangenen Moskitos eine Infektion aufweisen. Eine angemessene Datenanalyse sollte daher Modelle anwenden, welche raum-zeitliche Korrelation und den Überschuss an Nullwerten berücksichtigen, um Über- oder Unterschätzung der Parameter zu vermeiden. Zero-

inflated geostatistische Modelle berücksichtigen raum-zeitliche Korrelation, indem gebietsspezifische und Zeitintervall abhängige random effects eingeführt werden, wodurch die Anzahl der zu schätzenden Parameter steigt. Bayessche Modelle, implementiert durch Markov chain Monte Carlo (MCMC), ermöglichen die Anpassung von hoch-parametrisierten Modellen.

In dieser Arbeit werden zero-inflated Bayessche Modelle angewendet, um die durch Malaria bedingte Mortalität in allen Altersgruppen mittels großer, korrelierter und sparse (dünnbesetzt) Datensätzen, welche vom Navrongo und Manhica HDSS zwischen 2001 und 2004 gesammelt wurde, zu schätzen. Die Beiträge dieser Arbeit waren (i) die Beschreibung der HDSS Datenmerkmale und relevanten Analysemethoden; (ii) die räumlich-explizite Schätzungen der Intensität der Malariaübertragungen in monatlichen Intervallen; und (iii) das Verhältnis zwischen Gesamtmortalität und der Malariaübertragungsintensität in allen Altersgruppen.

Kapitel 2 beschreibt die Merkmale der MTIMBA Daten. Jene sind große geostatistische, zeitliche, saisonale und zero-inflated Daten. Die Mortalitäts- und Moskito-Daten waren nicht angeglichen, da sie in unterschiedlichen Gebieten und Zeitperioden erfasst wurden. Zero-inflated Bayessche zeitlich-räumliche Modelle sind hinsichtlich der Analyse solcher Daten der neueste Stand der Technik. Der exakte statistische Prozess wurde durch Modellierung der Sporozoitenrate (SR) Daten des Manhica HDSS aufgezeigt. Die Analyse der MTIMBA Daten wurde genutzt als ein Weg um SSA Kapazitäten durch Kursarbeiten, Seminare und Mentorschaft aufzubauen. Die durch das Projekt generierten Daten sind relevant zur Beurteilung der Heterogenität der Malariaübertragung innerhalb und zwischen Gebieten.

Die Navrongo Malaria Expositionsabbildungen, welche in Kapitel 3 beschrieben wurden, basieren auf zero-inflated geostatistischen Modellen. Diese wurden separat auf die binomialen

SR Daten und die negativ binomialen Zählungsdaten der Moskitoarten angewandt. Zusätzlich zu Kovariaten bzgl. Klima, Umgebung und Saisonalität beinhalteten alle Modelle räumliche sowie zeitliche Korrelation. Die Schätzungen der entomologischen Impfungsrate (EIR) wurde als Produkt der geschätzten Bissrate und der SR hergeleitet. Die beobachtete EIR in diesem Distrikt war >100 infektiöse Bisse/Person/Jahr. Distanz zu Gewässer, Tagestemperatur und Vegetation waren die Hauptprädiktoren der Moskitodichte der zwei Spezies. Die EIR Karten zeigen eindeutig auf, dass die zeitliche Heterogenität stärker war als die räumliche Variation in diesem Gebiet. Gleiche Ergebnisse ergab die Analyse der zwei MTIMBA Gebiete Rufiji (Tansania) und Kisumu (Kenia).

Monatliche Abbildungen der Malariaexposition (Kapitel 3) wurden verknüpft mit den nächstgelegenen Gebieten, in denen Mortalität beobachtet wurde (siehe Kapitel 4). Daten bzgl. des Todeszeitpunkts wurden in monatliche Intervalle eingeteilt um Bernoulli und binomiale Daten zu generieren, welche mittels logistischen Regression modelliert wurden. Räumlich-zeitliche Modelle wurden angepasst um das altersspezifische Mortalitätsrisiko zu schätzen. Das Model umfasste zwei Kovariaten – log-transformierte EIR Schätzungen mit ihren Messabweichungen und Alter. Die ITN Variable war nur enthalten in den Modellen für Neugeborene, Postneonatale (1-11 Monate) und Kinder. Die Analyse zeigte eine positive log-lineare Beziehung zwischen Gesamtmortalität und Malariaexposition in allen Altersgruppen >=60 Jahre. Der Gebrauch von ITN zeigte einen schützenden Effekt bei allen Kindern unter fünf Jahre. Dies bestätigt die Ergebnisse aus der Analyse der Rufiji und Kisumu HDSS.

Die Methode, welche in Kapitel 3 und 4 zur Schätzung der Abbildungen der Malariaexposition und des Mortalitätsrisikos angewandt wurden, wurden erweitert um die Manhica HDSS

(Mosambik) Daten zu analysieren und die Relation zwischen Mortalität-Malaria Übertragung in dieser Region zu beschreiben (Kapitel 5). Die räumlich-zeitlichen altersspezifischen Modelle umfassten die EIR Schätzungen mit ihren Messabweichungen (um Unsicherheit der Vorhersage zu berücksichtigen) und Alter als Kovariaten.

Die Distanz zum nächsten Gewässer war der einzige wichtige gemeinsame Prädiktor für *An. funestut* und *An. gambiae* Moskitodichte. Die Intensität der Malariaübertragung ist in diesem Gebiet beständig zurückgegangen. Die model-basierten Ergebnisse zeigen eine positive log-lineare Relation zwischen Gesamtmortalität und Malariaexposition in allen Altersgruppen auf (Neugeborene (0-28 Tage), Postneonatale (1-11 Monate), Kinder (1-4 Jahre), junge Menschen (5-14 Jahre), Erwachsene (15-59 Jahre) und alte Menschen (>=60 Jahre).

Diese Arbeit trägt zu weiterem Wissen über die Malaria-Mortalität Beziehung bei. Eine positive Assoziation zwischen Mortalität und Malariaexposition bei Kindern unter fünf Jahre stimmt mit den Ergebnissen der MTIMBA Gebieten Rufiji und Kisumu überein. Vervollständigung der Analysen in den verbleibenden Gebieten und eine anschließende Meta-Analyse werden einen großen Beitrag zur Malaria-Epidemiologie darstellen. Zukünftige Arbeit sollte eine Kohortenstudie berücksichtigen, um festzustellen, ob Malariakontrollinterventionen eine Verlagerung des Alters bzgl. der erworbenen Immunität verursacht haben.

Acknowledgements

There are many individuals and institutions I would like to express my sincere thanks for their contributions towards my PhD training. I may not be in a position to list all the names but I do really appreciate your contributions. I am very grateful to my supervisor PD Dr. Penelope Vounatsou for all the support you gave me throughout this journey. I do sincerely appreciate your guidance, mentorship, patience and encouragement. Thank you for accepting me on this PhD program that has exposed me to spatial epidemiology. I have really learnt a lot from you.

To Prof. Thomas Smith, I am so grateful for all the scientific support. I do appreciate all the constructive comments, expert reviews and for accepting meetings even on short notice. Special thanks also go to Prof. John J. Aponte for accepting to be my co-referee.

I would like to thank the management of the Swiss Tropical and Public Health Institute (Swiss TPH) for creating a good working environmental for people with different cultural backgrounds. Prof. Dr. Marcel Tanner, the Director and "my Faculty representative", thanks for all the support. Many thanks also go to Prof. Mitchell Weiss, Prof. Nino Künzli and other members for steering the EPH department. I appreciate the support from Christine Mensch, Margrit Slaoui, Christine Walliser, Maya Zwygart, Zsuzsanna Gyorffy, Dagmar Batra, Beatrice Wackerlin, Doris Stamm and the entire administrative team. I must say that you are really a special team. I am also grateful to the IT and library teams, thanks for responding to my emails and phone calls. To members of the Swiss TPH academic team, I like your style of teaching.

I wish to express my sincere thanks to the members of Biostatistics unit (Modelling and Mapping group) both present and previous namely; Laura Goşoniu, Konstantina Boutsika, Ronaldo

Scholte, Nadine Schur, Dominic Goşoniu, Federica Giardina, Venena Jürgens, Frédérique Chammartin, Alex Karagiannis and Eveline Hürlimann for all the support. I do appreciate the solidity. Thank you for allowing me to run my models on your computers. Special thanks go to Venena Jürgens for translating the summary of this thesis into German.

The MTIMBA team; Nyaguara Ombek Amek, Susan Rumisha, Bernadette Huho and Eric Diboulo, thanks for the team work and the brotherhood.

Special thanks go to all members of the Swiss TPH soccer team, I like your company. To all my African brothers and sisters from Kenya, Tanzania, Ghana, Burkina Faso, South Africa, Nigeria, Ghana, Côte d'Ivoire and other countries not listed, thank you for organizing the social gatherings. I cannot forget the Christmas and Easter days we celebrated together. I commend my brother Suleiman Jembe for uniting Africans in Basel.

Back to Uganda, I do sincerely thank the Makerere University administration for giving an opportunity to go back to school. Special thanks go to the management and entire staff of Makerere University School of Public Health for the support you gave me whenever I needed it. I thank the Chair and all the members of Epidemiology and Biostatistics department for taking an extra task while I was away. I am sincerely thankful to my colleagues, Drs. Nazarius M. Tumwesigye, Lynn Atuyambe, John Ssempebwa and others not listed for all the support I received from you.

My Mentor, Prof. Fred Wabwire-Mangen, thanks for all the support. I appreciate the phone calls, email messages and all the words of encouragement.

My good friend Dr. David O Mukanga, I do really appreciate your help. The discussions we held when you were writing your PhD thesis helped me also to sharpen my work. I cannot forget your words of encouragement.

Dr. Richard Muwonge, my very good old friend and brother, thanks for the evening phone calls. I am very grateful for all the advice, encouragement and also being ready to listen to me all the time. You really helped me a lot in this struggle.

I am extremely indebted to my family; Mum & Dad (Mr. & Mrs. Jerome Ddamba of Bujumba-Butambala), my siblings, in-laws, Kojja Edward Kalanzi-Ssewannyana, Rev. Fr. John L.Ssenkooto and others for the support, love and encouragement. I thank my parents for understanding the value of education.

My dear Sarah, I thank you for the love, support, patience and for standing in my shoes all the time I have been away. It has not been an easy task but God has finally seen us through. I thank the family of Mr. & Mrs. Edward Nkangi for supporting my people whenever I travel. May the almighty God reward you abundantly.

I thank all the MTIMBA investigators; Drs. Ricardo Thompson, Seth Owusu-Agyei, Salim Abdullah and Diadier Diallo for conceptualizing this project. I sincerely thank the management of Manhiça and Navrongo HDSS for giving us access to the data.

This work was financially supported by the Kanton of Basel Education Department, the Swiss National Science Foundation (project Nr. 325200-118379) and the Swiss-South Africa Joint Research Program (project Nr. JRP IZLSZ3_122926). Data collection was carried out under Tropical Disease Research (WHO-TDR) grant.

Chapter 1: Introduction

1.1 Background

Plasmodium falciparum malaria infection is one of the major causes of morbidity and mortality. In 2010, approximately 2.5 billion people globally lived in the regions that were exposed to *P. falciparum* (Gething et al., 2011). There were an estimated 216 million episodes of malaria world wide in 2011 and 81% of them occurred in Africa. Ninety one percent of the cases were due to *P. falciparum* (Cibulskis et al., 2011; Hay et al., 2010; WHO, 2011). More than half a million estimated deaths in 2011 were attributed to malaria. Most of the deaths occurred in Sub-Saharan Africa (81%) especially among children under five years of age. Figure 1.1 shows the global distribution of malaria risk. Malaria is present in 106 countries mainly in the tropics and subtropical regions.

Although there is a reported global decline in both morbidity and mortality, the figures still show a huge burden on the sub-Saharan Africa (SSA) (Murray et al., 2012; WHO, 2011). Reductions in malaria cases and deaths have been attributed to scaling up of the World Health Organization (WHO) recommended interventions namely; insect treated nets, indoor residual spraying (IRS), intermittent preventive treatment (IPT) during pregnancy, parasitological confirmations using either microscopy or rapid diagnostic tests (RDT) and treating all confirmed malaria cases with artemisinin combination therapy (ACT) (WHO, 2011). All these initiatives aim at reducing malaria infection in humans. However, reducing malaria exposure in endemic countries is likely to shift the age of acquired immunity leading to cases of severe disease in older children (Snow and Marsh 1995).

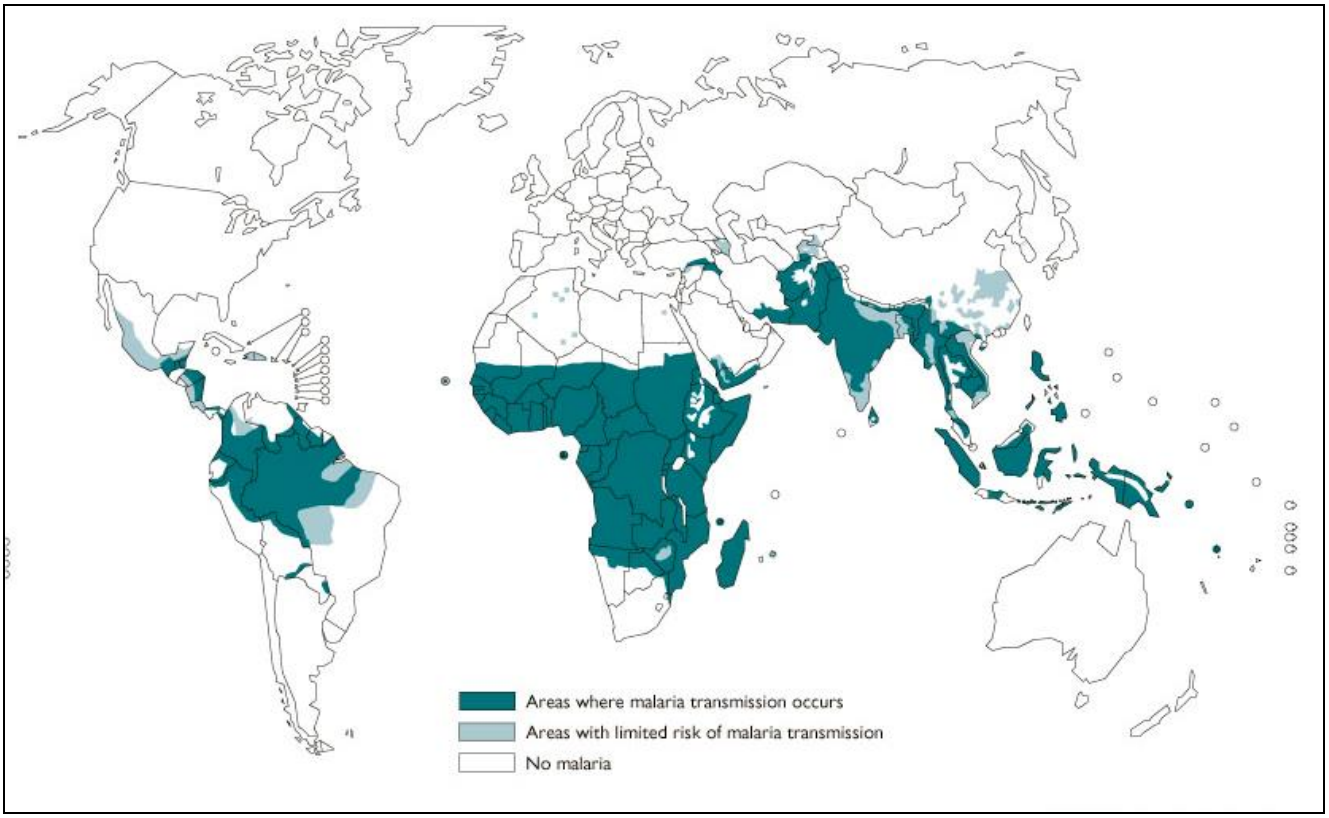


Figure 1. 1: Global distribution of malaria risk

Source :(http://www.who.int/gho/map_gallery/en/: Accessed 17/5/2012)

1.2 Malaria transmission

There are four main plasmodia species that cause malaria in humans namely; *Plasmodium falciparum, P. malariae, P.ovale and P. vivax.* The parasite develops in two phases; the asexual within the human host and sexual taking place within the mosquito (Beier, 1998). *P. falciparum* is the most common species in the tropics including SSA where the disease has overburdened the region. *P. malariae* occurs alongside with *P. falciparum* in the tropics and sub-tropical countries. *P.ovale* is primary found in SSA, while *P. vivax* is distributed within tropical and temperate regions though rare in Africa (Rogerson and Carter, 2008). In recent years, human cases of

malaria have occurred with *Plasmodium Knowlesi* a monkey parasite mainly in south-eastern Asia (Collins, 2012).

The malaria parasite has a complex cycle that involves a definitive host (mosquito) and human. Transmission of the parasite is from human to human through bites from a female anopheles mosquito as indicated in figure 1.2. Infection of human host starts when a mosquito injects malaria parasites (sporozoites) into the blood. The sporozoites then travel to the liver cells where they multiply asexually. Liver schizonts become mature and rupture, releasing merozoites into the blood stream. The merozoites then invade erythrocytes after their release and evolve into ring forms called trophozoites, which in turn form schizonts where new merozoites develop and are released into the blood circulation after. The simultaneous waves of merozoites escaping and infecting more red blood cells result into symptomatic malaria disease. Part of the merozoites develops into male and female gametocytes after going into a couple of schizogonic cycles. When a mosquito bites an infected human, it ingests the gametocytes, which further mature into male or female gametes and sexual replication takes place producing zygotes. These zygotes develop into mature oocyst which bursts to release sporozoites that invade the salivary grand of the mosquito, thus completing the cycle (stages 1-6 in figure 1.2). The life cycles of all human plasmodia species are similar but only vary in the length of time taken to complete a particular phase.

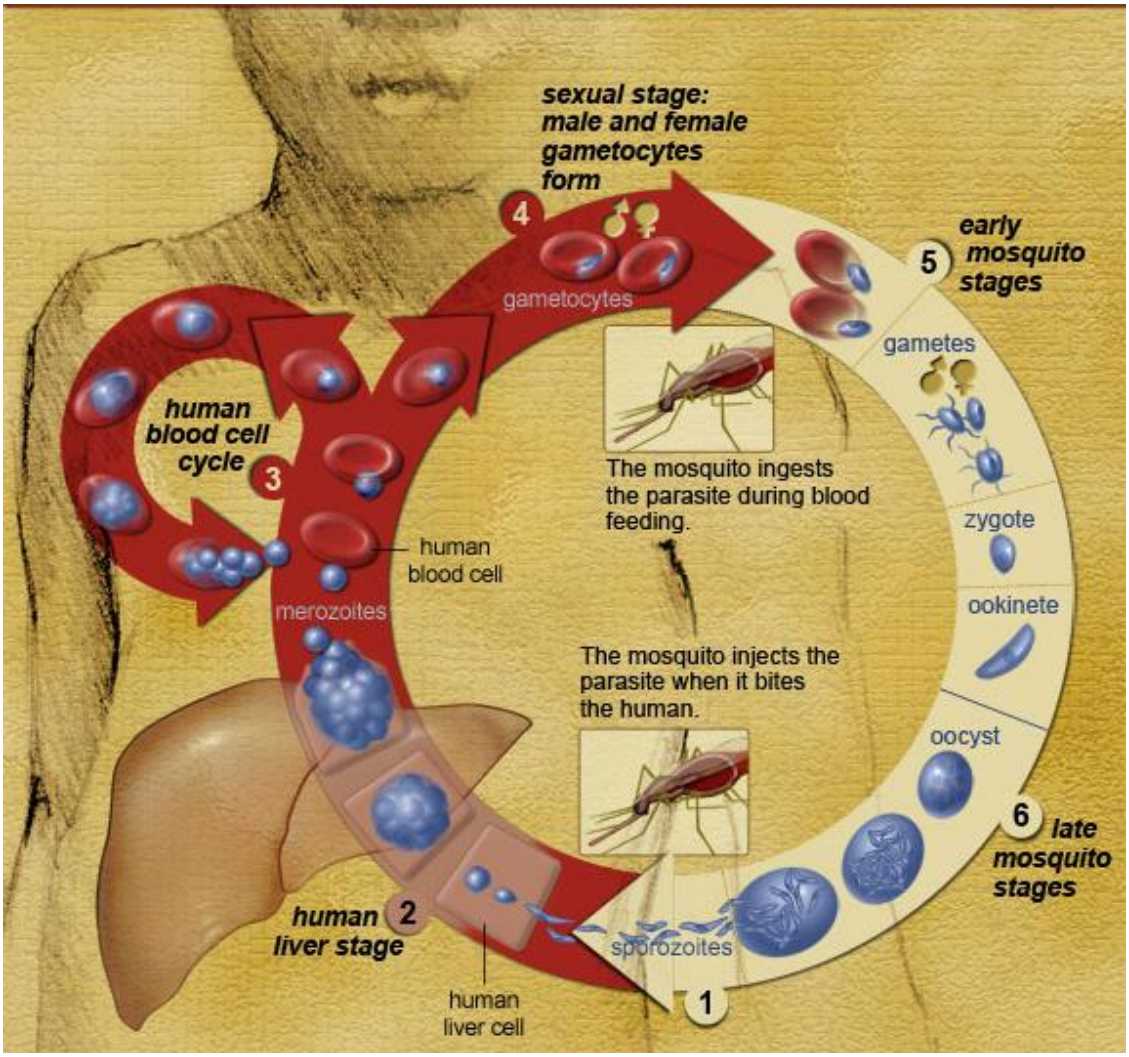


Figure 1. 2: Malaria transmission cycle

(Source: http://www.niaid.nih.gov/topics/malaria/pages/lifecycle.aspx : 10/5/2012)

Malaria vectors consist of various anopheles species with unique behaviours associated with ecological factors (Lindsay et al., 1998). Approximately 400 anopheles mosquitoes have been identified of which 30-50 species have the potential to transmit malaria to humans (Harbach, 2004). In SSA, malaria transmission is mainly driven by anopheles mosquitoes belonging to *An. gambiae* and *An. funestus* groups (Coetzee and Fontenille, 2004; Coetzee et al., 2000; Gillies and Mielion, 1968). The two species are mostly attracted to humans instead of other abundant

animals (anthropophilic). *An. funestus* breed in clear permanent fresh waters such as rivers and surround swamps while *An. gambiae* larvae have been found in temporal and shallow waters. The two salt-water sub-species of *An. gambiae complex* namely; *An. merus* and *An. melas* are found along the East and West African coasts respectively. Although the two species rest mainly indoor (endophilic), their feeding times differ (Gillies and Mielion, 1968).

1.2.1 Factors associated with malaria transmission

Climatic factors such as temperature, rainfall and humidity influence the mosquito abundance and disease transmission. Temperature is directly related to elevation of an area. Temperature decreases as altitude increases and consequently mosquito population, vector species and transmission intensity also change with elevation (Drakeley et al., 2005; Kristan et al., 2008; Shililu et al., 1998). Low temperatures are associated with prolonged larval development leading to increased mortality rates and hence low mosquito density (Minakawa et al., 2002). Temperatures above 22^{o}C have been considered to favour stable malaria transmission, however those above 32^{o}C cause high mosquito population turn-over, weak individuals and subsequently high mortality (Craig et al., 1999).

There is also a positive correlation between malaria disease and precipitation (Briët et al., 2008). Malaria transmission pattern follows rainfall distribution. Mosquito population increases in the middle of the rain season and reach a peak in the early part of the dry season. In equatorial region where two rainfall peaks are experienced and permanent swamps exist, fluctuation in the number of mosquitoes are much less than the Savannah area with single rainfall season (Gillies and Mielion, 1968). However, in areas with no holding swamps, heavy rains wash away mosquito larvae sites which reduce mosquito population and transmission. Conversely, end of the rain

season creates water ponds which act as favourable mosquito breeding sites even in the dry seasons. Malaria transmission therefore seems to be driven by climatic and ecological factors.

1.2.2 Measures of malaria transmission

Malaria indices are crucial in determining the burden on the population and also for measuring the progress towards control efforts. Clinical examination is one of the first methods used to quantify malaria endemicity (spleen rate) in a population (Baker et al., 1868). It involves determining the proportion of sampled population with enlarged palpable spleen at a particular time. Spleen rates (SPR) have been used to categorize areas according endemicity levels using children aged 2 to 9 years as hypoendemic (SPR: 0-10%), mesoendemic (SPR 11-50%) , hyperendemic (SPR : 50 – 75%) and holoendemic (SPR: >75%) (Kevin Baird et al., 2002).

Parasite prevalence is also another malaria index that is used to monitor endemicity in a population. It refers to the total number of people (new and old cases) with a positive blood smear test of the total number screened at a particular time point. Using passive surveillance approach, the burden of malaria can also be measured by considering all reported malaria cases over the total number of people seeking treatment in that particular health facility. However, such a method is challenged by poor record keeping in SSA and also low utilization of health facilities. Periodic malaria indicator surveys carried out in Africa can also act as good sources of such information.

Another parasitological measure of malaria risk is the clinical incidence, which refers to the number of new cases within a given time period. Annual parasite incidence is one of form of incidence countries usually use to compare malaria risk between communities, districts or countries (Hay et al., 2008; Kevin Baird et al., 2002). Although rapid diagnostic tests are currently used in malaria diagnosis even at community level (Mukanga et al., 2012; Murray et

al., 2008), systems for gathering all confirmed case data are absent in most countries in sub-Saharan Africa.

Serological tools (Drakeley et al., 2005) have been proposed to be used especially in countries that are tending to elimination where transmission intensity has gone down or where transmission is very low. Indices generated under this approach can be classified under either prevalence or incidence.

The entomological inoculation rate (EIR) , which is referred to as the number of infective mosquito bites received per person per unit of time is the recommended direct method for measuring transmission intensity in endemic areas (Beier et al., 1999; The malERA Group, 2011). It is derived as a product of the proportion of mosquitoes with sporozoites in their salivary glands (sporozoite rate) and human bite landing. The latter is measured by the number of mosquitoes trying to feed on an individual. Although the gold standard method for estimating EIR is human landing catches (HLC), this approach is considered unethical, time-consuming, labour intensive and expensive. Mosquitoes are instead captured using pyrethrum spray catches, exit trap catches and CDC light traps methods (Shaukat et al., 2010).

1.2 Malaria control interventions

The development and use of residual insecticides like dichlorodiphenyltrichloroethane (DDT) became prominent at the end of the Second World War in the fight against malaria. Malaria control strategies applied DDT to reduce the mosquito population while infected people were treated with quinine which was one of the available anti-malaria drugs (Stapleton, 2009). In the second half of the 20th century, indoor residual spaying (IRS) with DDT led to a substantial decline in malaria in Sri Lanka, the former Soviet Union and India. The successful malaria eradication pilot project was not extended to many other area due to high program costs,

emergence of resistance to DDT and community resistance to repeated house spaying (Arrow et al., 2004; Greenwood and Mutabingwa, 2002). In 1969, the malaria eradication strategy was formally abandoned at the 22nd World Health Assembly and a call for new malaria control strategies was made (Muturi et al., 2008).

Strategies such as primary health care (PHC) which involved community health workers in health service delivery were also adopted for malaria control in SSA (Christopher et al., 2011). A trial involving ITN and chemoprophylaxis that was carried out in a village-based PHC scheme in rural Gambia attributed reductions in mortality in children to treated nets (Alonso et al., 1991). The results prompted more funding from WHO for four trials in Gambia (D'Alessandro et al., 1995), Kenya (Nevill et al., 1996), Ghana (Binka et al., 1996) and Burkina Fuso (Habluetzel et al., 1997). The four clustered randomized trials reported protective efficacy of ITN among children. Due to observed benefits, more trials were further extended to other areas (Arrow et al., 2004).

In 1998, WHO established Roll Back Malaria initiative (RBM) with aim of reducing malaria mortality in endemic areas using ITN as one of the tools. The African heads of state summit on malaria held in Abuja, Nigeria in 2000 endorsed the initiative (Greenwood and Mutabingwa, 2002; Yamey, 2000). Development partners have availed funds to RBM to fight malaria burden in SSA using effective preventive and treatment methods (WHO, 2011). However, continuous monitoring of RBM indicators has been challenged by lack of reliable data caused by weak health systems (Greenwood and Mutabingwa, 2002).

1.4 Malaria related mortality

Persistent high mortality estimates have been reported in the malaria endemic area of the SSA with high concentration among children under five years (Lopez et al., 2006; Rowe et al., 2006; Snow et al., 1999). Although recent indicators show a decline in mortality rates in the region that have been associated with millennium goals interventions (Murray et al., 2012; Rajaratnam et al., 2010; WHO, 2011), the true picture might be different because most deaths in developing countries take place outside health facilities and are never recorded. Countries have applied strategies namely; use of insect treated nets and prompt treatment of malaria which have been associated with a reduction to mortality especially among the children (Binka et al., 1996; D'Alessandro et al., 1995; Lengeler, 2004; Phillips-Howard et al., 2003). SSA experience variation in malaria exposure, age pattern for acquired immunity and access to health services that have been associated with mortality (Hay et al., 2000; Kelly-Hope and McKenzie, 2009). It has been noted that interventions targeting reductions in malaria exposure to people in endemic areas are likely to increase the age for acquiring functional immunity. Severe malaria cases in high transmission areas reduce with increasing age as a result of early acquired *P. falciparum* immunity (Snow and Marsh, 1995; Snow et al., 1997). Proper implementation of malaria control activities requires also a clear understanding of how mortality relates to transmission intensity. Previous efforts to assess the malaria attributed mortality have been hampered by lack of reliable data which is caused by inefficient health systems in SSA. Snow et al,(Snow and Marsh, 1995) carried out a meta-analysis using previous studies from Africa and found no relationship between mortality and malaria transmission. Subsequent meta-analyses indicated a positive relation between mortality and transmission intensity among the infants but not in children (12-59 months) (Lengeler et al., 1997; Ross and Smith, 2006; Smith et al., 2001).

1.5 The MTIMBA project

The United Nations millennium goals targeting malaria focus on scaling up of sustainable preventive and therapeutic interventions in countries that are overburdened by the disease. Measuring impact of such interventions requires reliable sources of data that are not available in most of these countries (Mathers et al., 2005). Many countries are now relying on health and demographic surveillance systems (HDSS) that were set up to routinely collect demographic and health related outcomes data within a defined geographical area to measure the effect of various interventions (Ngom et al., 2001; Tollman and Zwi, 2000).

The INDEPTH, a network of HDSS in developing countries established the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project in early 2000 to generate data that will provide evidence about malaria control efforts in SSA. The project aimed at assessing the levels of malaria transmission intensity; establishing the relationship between all-cause plus malaria mortality and malaria transmissions intensity taking into account the effect of disease control interventions. The project was linked into the routine activities of HDSS and field work was carried out for a period of three years. There are currently 19 countries with HDSS in the INDEPTH network and 12 are found in Africa with 26 sites as shown in Figure 1.3.

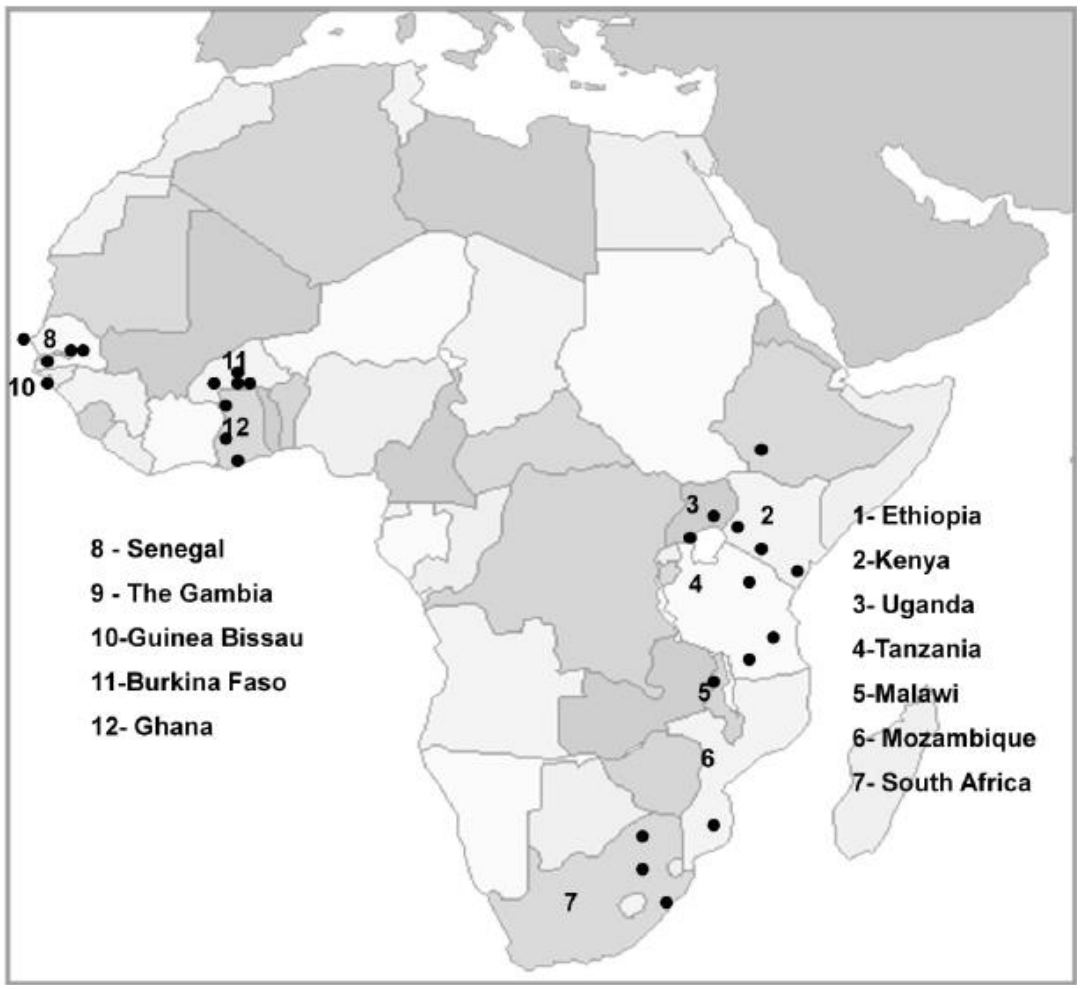


Figure 1. 3: Countries with Health and demographic systems in Africa

Six HDSS within five countries namely; Burkina Faso, Ghana, Kenya, Mozambique and Tanzania provided comprehensive data for the project for 3 years (2001 to 2004). Mortality data were obtained from the sites' databases of continuously monitored of demographic events. Entomological data across sites were collected using CDC light traps in order to obtain unbiased and comparable EIR estimates.

The MTIMBA-HDSS data were collected at large number of fixed compounds that are close to each other over the project period. Such geostatistical data are correlated in space because compounds close to each other share similar exposures.

The entomological data are correlated in time because they were collected fortnight over a three year period. Malaria transmission on the other hand is influence by climatic and environmental factors. Malaria transmission intensity tends to follow a climatic pattern of the area. There were many compounds in the dry season with zero mosquito catches. Similarly, the number of compounds with zero catches reduced in the wet seasons relative to the dry season. The influence of ecological and climatic factors leads to sparse entomological data.

The mortality and entomological data were not directly obtained from the same compound. Mortality was monitored in the entire HDSS while entomological data were collected from randomly selected compounds. Such data are known to be misaligned in space and time. To align the data, we need to develop predictive models that will estimate malaria transmission intensity at unsampled locations.

1.6 Modelling malaria spatial temporal heterogeneity

Advances in Geographical Information Systems (GIS) have enabled accurate geocoding of locations where data are collected. This has led to formulation of spatio-temporal databases in many fields including malaria hence promoting spatial data analysis. Proximity in space and time introduces spatial and temporal correlations (Cressie, 1993). Standard statistical models assume independence of observations. Ignoring spatio-temporal correlation may result into under or over-estimation of the significance of model covariates.

In malaria epidemiology, space and time heterogeneity can be modelled by Bayesian geostatistical models in order to obtain posterior distributions of EIR indices for small areas and time periods. These models relate entomological data to environmental factors after taking into account spatial and temporal correlation (Cressie, 1993). Recently geostatistical models have been used to assess malaria risk mostly from parasitological surveys (Kazembe et al. 2006; Noor

et al. 2009; Gosoniu et al. 2010; Riedel et al. 2010; Gething et al. 2011; Gosoniu et al. 2011; Giardina et al. 2012).

The MTIMBA project collected entomological data to estimate EIR, a recommended measure of transmission intensity in endemic areas. Rigorous analyses of these data therefore should take into account data characteristics namely; the distribution, collections over large number of geo-referenced compounds, spatio-temporal correlation, seasonality and misalignment in order to reduce bias in parameter estimation.

The geostatistical models for entomological data are either binomial (sporozoite rates) or Poisson/negative binomial (density) with additional parameters at each household location. The large number of households monitored in HDSS increases the number of parameters to estimate. The spatial dependence in each model is accounted for by introducing location-specific random effects which are assumed to be latent observations derived from multivariate spatial process with a zero mean. The covariance of the spatial process assumes a correlation function of distance between any pair of locations. The time correlation can also be modelled by introducing temporal random effects at defined time points (weekly, bi-weekly or monthly). This creates highly parameterized geostatistical models which makes maximum likelihood inference unstable. Bayesian models implemented via Markov chain Monte Carlo simulation (MCMC) addresses fit of highly parameterized models (Gelfand and Smith, 1990). However, with large number of locations (N>1000), geostatistical computation involves matrix calculations such as inverses and determinants that become very slow and probably infeasible. This computational challenge is informally referred to as "the big N problem" (Banerjee et al., 2003). Different approaches to tackle the large N problem have been proposed but have not fully removed the computational difficulty. These include use of low rank splines (Lin et al., 2000) and kernel convolutions

(Paciorek and Schervish, 2006) where a spatial process is approximated using a lower dimension subspace. Rue et al.(Rue and Tjelmeland, 2002) proposed approximating the spatial process by a Markov random field, a method that is suitable for locations on a regular grid. However the above methods have not been applied to health data where locations are clustered. In addition, these approaches are not straight forward to implement.

Banerjee et al. (Banerjee et al., 2008) and Finley et al.(Finley et al., 2009) proposed a much easier approach of a predictive process approximation which estimates the spatial process from a subset of locations (knots) with a reduced dimension ($N^{*} < N$). This approach improves the computational speed since the matrix to be inverted reduces according to the sample size of the knots (N^{*} by N^{*}). However, selection of knots with all the characteristics of the original space is a challenge. The team proposed selection of knots where the spatially averaged prediction variance (SAPV) is minimized (Finley et al., 2009). In addition to SAPV, Gosoniu et al. (in press) compared other sampling methods in selecting knots namely; balanced sampling (Deville and Tillé, 2004) and minimax space filling (Johnson et al., 1990) in order to estimate the computational costs. Findings indicate that models performed different when the selected number of knows is small (<200). Large "N" is still an on-going research topic in statistics and therefore the relevant softwares are not readily available. Currently available softwares such as BayesX (Brezger et al., 2005) and spBayes (Finley et al., 2007) are still under development.

Lack of standard software to analyse large geostatistical data generated by the MTIMBA project delayed the entire process of answering the project's research question. In addition, longitudinal entomological data are characterized with large number of locations with zero (zero-inflated).

1.7 Thesis Objectives

The aim of this work was to estimate malaria attributable mortality across all age-groups using large correlated data collected from health and demographic surveillance sites in west and southern Africa.

1.7.1 Specific objectives

- To describe the MTIMBA project data and identify relevant statistical issues.
- To estimate malaria transmission intensity in Navrongo and Manhiça HDSS.
- To relate all-cause mortality to malaria exposure using data collected from Navrongo and Manhiça HDSS.

1.8 Structure of the thesis

This thesis is organized as follows. Chapter 2 describes the MTIMBA project data characteristics and associated statistical issues. Chapter 3 presents an application of zero-inflated Bayesian geostatistical models to estimate monthly malaria exposure surfaces for the Navrongo HDSS, Ghana. In chapter 4 all-cause mortality was related to EIR estimates generated in the previous chapter. Chapter five presents the effect of malaria transmission intensity on mortality in Manhiça HDSS, Mozambique. A concluding discussion including the overall conclusion and study limitations are given in chapter 6.

Chapter 2: Malaria transmission intensity and mortality burden across Africa project; statistical issues and approaches to data analysis

Authors: Kasasa S[1, 2,3,], Rumisha SF[1,2,4], Amek N[1,2,5], Diboulo E[1,2,6], Huho BJ[1,2,7], Thompson R, Abdullah S[7], Diallo D, Owusu-Agyei S[8], Smith T[1,2], Vounatsou P[1,2*]

[1]Swiss Tropical and Public Health Institute, Socinstrasse 57, P.O. Box, 4002 Basel, Switzerland
[2]University of Basel, Switzerland
[3] Makerere University School of Public Health, College of Health Sciences, Kampala, Uganda

[4]National Institute for Medical Research, Dar es Salaam, Tanzania,
[5] Kenya Medical Research Institute/Centers for Disease Control and Prevention, Kisumu, Kenya
[6]Nouna, Borkina Faso

[7]Ifakara Health Institute, Dar es Salaam, Tanzania,
[8]Kintampo Health Research centre, Ghana Health Services, Ministry of Health

*Corresponding Author

This manuscript is prepared for submission to Malaria Journal

Abstract

The relationship between mortality and malaria transmission intensity remains unclear but mortality data are routinely collected in Health and Demographic Surveillance Sites (HDSS) many of which are in malaria-endemic areas. To study the relationship of mortality with malaria transmission, the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project collected entomological data over a 3 year period, from a very large number of locations within 8 HDSS. Given the small number of sites, between-HDSS analysis is not very informative. The within-site variation contains considerable information about the mortality-malaria relationship but analyses of this need to account simultaneously for the large number of locations sampled, the spatio-temporal correlation, seasonality, and the sparsity of the data, with large proportions of zero values. The mortality and entomological data are also misaligned because they were collected at different locations and time points. This means that the optimal analytical approaches require non-standard methods. In this paper, we described data features and statistical issues of the MTIMBA data, propose data-driven Bayesian methods for their analysis and provide the current status of the project. The methods are illustrated by the modelling sporozoite rate data from the Manhiça DSS.

Key words: INDEPTH; Spatio-temporal analysis; Bayesian inference; zero-inflated models; MTIMBA; Malaria transmission

2.1 Introduction

Evaluating the effectiveness of health interventions require appropriate data on morbidity, mortality and their specific causes in order to derive trends over time. However, recording vital events such as birth, death and migration in most African countries is still inadequate (Mathers et al., 2005). This is partly due to the fact that most births and deaths occur in homes and are never reported in national statistics. Countries rely on information generated from censuses and surveys which are not continuously carried out. Lack of vital registrations on population and health led to the establishment of health and demographic surveillance sites (HDSS) to collect routinely all related demographic and health outcomes within a defined geographical area(Ngom et al., 2001; Tollman and Zwi, 2000). Countries are currently using HDSS data for planning, policy formulation and monitoring disease outcome including malaria (Adazu et al., 2005; Byass et al., 2002; Deressa et al., 2007; O'Meara et al., 2008; Snow et al., 2004). In 1998, the International network of field sites with continuous demographic evaluation of populations and their health (INDEPTH) was set-up with an aim of improving population-based health information in developing countries (Ngom et al., 2001). Currently, there are 42 sites in the network within 19 countries where 69% are located in Sub-Saharan Africa (SSA).

Malaria is a common infectious disease transmitted by anopheles mosquitoes in the SSA countries where the majority of network sites are situated (Bryce et al., 2005; Morris et al., 2003; Rowe et al., 2006) . Transmission intensity especially in Sub-Saharan African is heterogeneous. It ranges between zero and more than 1000 infective bites per person per year (Beier et al., 1999). However, while severe malaria has a high case fatality rate, and substantial reductions in mortality have been observed in field trials insecticide-treated nets (Akachi and Atun, 2011; Eisele et al., 2010; Lengeler, 2004), the quantitative relationship between malaria transmission

intensity and mortality remains unclear(Gemperli et al., 2004; O'Meara et al., 2008; Ross and Smith, 2006; Smith et al., 2001; Snow and Marsh, 2002, 1995). A study that reviewed mortality and entomological inoculation rate (EIR) data from Africa documented a positive relationship between infant mortality and EIR (Smith et al., 2001). The positive association was not observed for children aged 12 to 59 months. Geostatistical analysis using the Mapping Malaria Risk in Africa (MARA) and the Demographic and Health Surveys (DHS) databases found no relationship between malaria risk and infant mortality (Gemperli et al., 2004).

In directing further interventions targeting Millennium Development Goals (MDG) on Malaria, the INDEPTH network established the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) initiative. The field work was implemented in 8 HDSS between 2001 and 2004 with the aim of examining the relationship between mortality and malaria transmission intensity taking into account interventions implemented by participating sites.

A standard protocol was developed for estimating the Entomological Inoculation Rate using CDC light traps to estimate indoor densities of host-seeking Anopheles, and ELISA assays to assess sporozoite-positivity in the vectors. The protocol also involved calibration of the CDC light traps against human landing collections in order to estimate the exposure of a sample of the human population, representative in space and time. These data could be linked to the data on both all-cause mortality, and cause-specific mortality derived from verbal autopsies.

Despite the large amount of data collected, the variation between sites in mortality rates did not show any clear relationship with estimates of malaria transmission intensity at the site level. Thus the key analyses consider mainly the variation within sites.

In this paper we describe the MTIMBA project, data challenges, appropriate methods for their analysis and current status. The methods are illustrated with the modelling of sporozoite rate data from the Manhiça DSS.

2.2 MTIMBA project

2.2.1 Project sites

Ten demographic surveillance sites in sub-Saharan Africa participated in the project but only 8 finally provided comprehensive mortality data, comprising Manhiça in southern Africa (Aranda et al., 2005); Rufiji, Ifakara and Kisumu in East Africa(de Savigny et al., 2004; Somi et al., 2007); Nouna, Naikhar, Oubritenga, Kourweogo and Navorongo in West Africa (Appawu et al., 2004a; Diallo et al., 2004; Etard et al., 2004; Hammer et al., 2006.; Konaté et al., 2011). Figure 1.3 shows countries (numbered 2, 4, 6, 8 11 and 12) where MTIMBA sites are located and their details of have also been described in the INDEPTH monograph (2000) and elsewhere (Ngom et al., 2001) .

2.2.2 Data collection

Entomological data

Mosquito collection was performed all the year around using light trap catches in rooms of randomly selected members of the HDSS population. The intention was that each month a minimum of 10 all-night light trap collections were conducted within each HDSS site. For logistical reasons, it was not possible to obtain a simple random sample of the host-seeking mosquitoes by locating the collections independently of each other. Instead, each site used a slightly different sampling strategy depending on available resources and local settlement

patterns, aiming to obtain an unbiased estimate of the numbers of biting mosquitoes. Sampling methods for Oubritenga, Kourweogo and Ifakara HDSS were different from the rest.

Traps were hung at about 1.5m above the floor next to the bed of the index person. Other people in the same room without bed nets were also provided with untreated nets for that specific night. Light traps were operated from sundown to sunrise from each of the randomly selected compounds. For calibration purposes, at least 30 human landing catches (HLC) were set over the transmission period in order to obtain a correction factor (Lines et al., 1991). HLC involved 2 individuals seated collecting mosquitoes landing on their exposed legs using a torchlight and test-tube or aspirator. Each pair of mosquito collectors worked for six hours per night. HLC fieldworkers were given malaria prophylaxis based on the country's treatment guidelines. All mosquitoes were transported from the field to the laboratory in a cool box or in a tube containing desiccant for further processing. Heads and thoraces of light trapped anopheles were tested for *Plasmodium falciparum* circumsporozoite protein *(CSP)* using enzyme linked immuno-sorbent assay (ELISA).

The mosquito density, and sporozoite data were used to compute Entomological Inoculation Rates, for specific locations and time periods, as the product of the estimated proportion of host-seeking mosquitoes that are sporozoite positive (sporozoite rate) and the estimated number of mosquitoes biting a mosquito collector in unit time (man biting rate) (Beier et al., 1999).

Mortality data

Mortality data were extracted from routinely collected HDSS databases. Cause- specific mortality data were obtained from a modified verbal autopsy (VA) tool derived from both, the World Health Organization (WHO) and the standard, site-specific VA questionnaires. Questionnaires were translated into local languages in order to suite the local socio-cultural

environment. VA interviews were conducted within the first three month of death. Interviews administered by trained researchers were analyzed by two independent physicians in order to determine the probable cause of death. Whenever consensus was not reached by the two, a third assessment was conducted on the same questionnaire. Cause-specific death was coded according to either International Classification of Diseases (ICD-10) or locally- derived systems.

Informed consent

At the initiation of the study, HDSS sites explained to their communities the project objectives and the approaches to use. Verbal consent was sought from household heads, index persons and other members in the room where traps were to be set. At each survey round, written consent was sought from fieldworkers who performed human landing catches. Details about study, anticipated hazards were given before recruitment. In addition, sites supplied HLC individuals malaria prophylaxis based on countries treatment guidelines [MTIMBA protocol, unpublished].

2.3 Data characteristics

These large amounts of data are spatially correlated because neighbouring locations share common exposures such as interventions, land use, climate and environmental factors. The longitudinal nature of such data also introduces a temporal correlation alongside the mortality data obtained from the HDSS sites during the project period. Seasonal changes in the weather, influence mosquito behaviours and malaria transmission(Abellana et al., 2008; Mabaso et al., 2007; Oesterholt et al., 2006; Okello et al., 2006). Seasonality and temporal trends are therefore present in the MTIMBA data. The entomological data collected over time usually contain mosquito collections with either no mosquitoes or zero infected mosquitoes. Over 50% of the

collections in all sites had no mosquitoes (Amek,. et al., 2012; Kasasa , et al.,2013; Rumisha, et al., 2012)**.** The large amounts of zeros cause data to be over-dispersed.

Statistical analysis

Appropriate data analysis should take into account spatial-temporal correlation in order to avoid over or underestimation of statistical significance for the covariates (Cressie, 1993). Bayesian geo-temporal statistical models are the state-of-art methods for analyzing the DSS mortality and MTIMBA entomological data. However, assessing the relationship between mortality and transmission has been delayed by the computational difficulties involved in the model fit. Modelling of these data takes into account spatial correlation by incorporating random effects at the observed locations. These are treated as latent variables arising from a spatial process quantified by a multivariate normal distribution (Gaussian spatial process)(Diggle et al., 1998) . Spatial correlation is taken into account in the covariance matrix of the process by assuming a correlation function of distance between any pair of locations. Model fit is complicated due to large numbers of parameters. Bayesian models implemented via Markov chain Monte Carlo simulation (MCMC) addresses fit of highly parameterized models however; geostatistical computation involves matrix calculation such as inverses and determinants. For large number of locations, these calculations are infeasible ("large N problems"). Banerjee et al(Banerjee et al., 2008b) and Finley et al.(Finley et al., 2009) proposed estimation of the spatial process from a sample of locations. Gosoniu GD et al. (Gosoniu D. et al., 2011) assessed different sampling schemes using DSS mortality data and concluded that balanced sampling (Deville and Tillé, 2004) and space filling algorithms (Johnson et al., 1990) provide a good sample tools in obtaining the sub-set of locations (knots). For the analysis of the MTIMBA data, Rumisha et al

(Rumisha, et al., 2013) proposed an approach of selecting the size and sub-sample by comparing the variogram of the full data with that of the sub-samples.

2.3.1 Seasonal and temporal data

.

Seasonality is often captured either as a binary covariate dry/wet or as a trigonometric function described as follows;

$f(t) = b_1 \cos\left(\frac{2\pi t}{T}\right) + b_2 \sin\left(\frac{2\pi t}{T}\right), \qquad t = 1, 2, ..., n$ where $f(t)$ is the seasonality function, T, the season length and b_1 plus b_2 are the components for amplitude and phase (Stolwijk et al., 1999). The function $f(t)$ is typically included in the regression model to account for seasonality in the data.

Temporal random effects (ε_t) can modelled via an autoregressive [AR (k)] stationery process of order k (Hay and Pettitt, 2001). The above trigonometric models have been applied for the analysis of Rufiji MTIMBA data (Rumisha et al., 2012) .

2.3.2 Sparse data

Standard regression approaches that ignore the excess zeros in data usually fail to provide adequate fit to the data (Ridout et al., 2001), so zero-inflated regression models were used to accounting for the excess zeros. Entomological data are characterized by large number of locations with either no mosquitoes or proportion with sporozoites in their glands. Zero-inflated models are formulated as two-component mixture models; one corresponding to the structural zeros due to unmeasured covariates and another one to the distribution that generated the data. Formally, the model is written as follows; $P(Y_{it} = y_{it}) = \begin{cases} \pi_{it} + (1-\pi_{it}) f(y_{it}), & y_{it} = 0 \\ (1-\pi_{it}) f(y_{it}), & y_{it} > 0 \end{cases}$ where,

$f(y_{it})$ is the standard distribution and π_{it} , is the probability (mixing proportion) of observing a structural zero. The mixing proportion is considered either as a constant or as a function of covariates depending on data fit (Lambert, 1992) . Such models have been applied mostly to count epidemiological data(Clements et al., 2006; Nobre et al., 2005; Soares Magalhães et al., 2011; Vounatsou et al., 2009), but sparse literature is available for binomial data(Hall, 2000).Bayesian zero-inflated binomial models have been developed by Amek et al (Amek, et al., 2011) for analysing the Kisumu MTIMBA data and applied by Kasasa et al (Kasasa , et al.,2013) and Rumisha et al(Rumisha, 2013) for the analysis of Navrongo and Rufiji data, respectively.

2.3.3 Misaligned data

Entomology and mortality data were collected within MTIMBA-DSS sites at different locations and over a time period making them time and spatially misaligned. Such data can be aligned by developing geostatistical models to predict the exposure at the outcome locations taking into account the prediction error as a measurement error in the covariate (Gemperli, 2003). For the MTIMBA-HDSS data, negative binomial (mosquito density) and binomial (SR) models by mosquito species were fitted separately. Bayesian kriging (Diggle et al., 1998) was applied to predict density and SR at the unsampled locations. EIR estimates were generated using model-based products of density and sporozoite rate at high resolution. Mortality from georeferenced compounds were then linked to the nearby EIR based on the minimum distance (Amek, 2013; Rumisha , 2013;Kasasa et al., In preparation).

2.3.4 Complementary data

Mortality may not only be influenced by malaria but also by ongoing interventions to both, the disease and the vector. Households' socio-economic status (SES) can be related to mortality as it determines choice and access to health care especially in rural communities (Rutebemberwa et al., 2009). Not all MTMBA sites have complete data for interventions and SES. Some have no data at all and others collected it before or after observing mortality in certain households. Mosquito breeding sites (i.e. swamps, ponds, seasonal rivers) and climatic factors affect malaria transmission in an area. Although these data are routinely collected by some sites, they were not included in the standard MTIMBA protocol. In addition not all compounds in the HDSS were georeferenced. Available weather and environmental data at very high spatial resolutions are too expensive. Remote sensing data at 250m to 1km spatial resolutions from Moderate Resolution Imaging Spectroradiometer (MODIS), African Data dissemination Service (ADDS) and HealthMapper can be used as potential proxies. These data are available at high temporal resolution (weekly or by weekly intervals). Lag time analysis or spatial variable selection procedures can be used to determine the period prior to data collection which climatic proxies can be based.

2.4 Statistical models

2.4.1 Modelling Sporozoite Rate data

Sporozoite rate are binomial data modelled via logistic regression. The number of positive mosquitoes $Y_{it}^{(S)}$ out of all tested (N_{it}) follows a binomial distribution; that is $Y_{it}^{(S)} \sim Bin\left(N_{it,} p_{it}\right)$ with parameter p_{it}, the sporozoite rate at location i and time t. These are

modelled together with covariates X_{it} , seasonality and random errors on the logit scale as, $\log it(p_{it}) = X_{it}^{T}\beta^{(S)} + \phi_i^{(S)} + \varepsilon_t^{(S)} + e_i^{(S)}$, where $\beta^{(S)} = (\beta_1, \beta_2, ..., \beta_k)^T$ is a vector of regression coefficients. The spatial random effects are assumed to originate from a Gaussian spatial process with zero mean and correlation matrix $R^{(S)}$, where $\phi_i^{(S)} = (\phi_1, \phi_2, ..., \phi_n)^T \sim N(0, \sigma_\phi^{(S)2} R^{(S)})$ and $\sigma_\phi^{(S)^2}$ is the spatial variance. A number of functions can be used to measure correlation between any particular pair of locations(Gelfand, 2007), the exponential one is frequently used, that is $R_{ij}^{(S)} = \exp\left(-d_{ij}^{(S)}\rho^{(S)}\right)$ where $d_{ij}^{(S)}$ is the Euclidean distance between locations i and j, and $\rho^{(S)}$ is the correlation decay parameter. Non spatial random effects $e_i^{(S)}$ with zero mean and variance $\sigma_e^{(S)^2}$ are added to the model to account for unexplained variability in the data.

2.4.2 Modelling mosquito density data

Mosquito densities are typically over-dispersed count data, best modelled by negative binomial distributions. Let $Y_{it}^{(D)}$ be the number of mosquitoes trapped at location i and time t. We can assume that $Y_{it}^{(D)} \sim NB(\mu_{it}, r)$, with μ_{it} and r corresponding to the mean and dispersion parameters respectively (Vounatsou et al., 2009a). The relationship between mean density of each species (μ_{it}) , the covariates X_{it} and the random effects is modelled as; $\log(\mu_{it}) = X_{it}^{T}\beta^{(D)} + \phi_i^{(D)} + \varepsilon_t^{(D)} + e_i^{(D)}$ where, X_{it} is the vector of covariates at location i for time t and $\beta^{(D)} = (\beta_1, \beta_2, ..., \beta_k)^T$, the vector of regression coefficients. All the random effects are defined and modelled in similar way as described in the sporozoite rate model above.

2.4.3 Modelling EIR

EIR data are treated as log-normally distributed (Gemperli et al., 2006; Himeidan et al., 2011) which leads to difficulties in modelling small area variation, because the data are often sparse, and often mosquito collections in the dry season capture no mosquitoes. However, EIR arises as a product of sporozoite rate and human biting rate derived from mosquito density. Sporozoite rate are binomial data while mosquito density count data follow either Poisson or negative binomial distribution. Proper statistical data analysis requires taking into account a distribution that generated that data. Applying independent logistic and negative binomial regression models to sporozoite rate and density data by mosquito species respectively would therefore lead to accurate EIR estimates. Zero-inflated analogues of the binomial (ZIB) and the negative binomial (ZINB) could be used to account for the effect of the large number of location traps with no infected mosquitoes. The two estimates, sporozoite rate and mosquito density are then multiplied together including a conversion factor for adjusting for light trap catches to man biting rate(Lines et al., 1991). At unsampled locations where entomological data are not available, Bayesian Kriging can be used to predict both sporozoite rate and density rate data (Diggle et al., 1998).

2.4.4 Modelling mortality and malaria transmission

The relationship between mortality and malaria is age-dependant since the disease morbidity is linked to immunity which develops with age. Therefore assessing such a relationship should also be studied at different age groups due to variations in mortality determinants (Becher et al., 2008a). Common age specific mortality categories include, neonatal (0-28 days), postnatal (1-11months), child (1-4 years), young people (5-14 years), adults (15-59 years) and old age (atleast 60 years). Survival models are appropriate in analyzing mortality data and they assume

continuous follow-up time (Cox, 1972). In the presence of time-dependant covariates, survival models are based estimated by logistic regression where discrete follow-up time is assumed(Allison, 1982; Singer and Willet, 1993). The occurrence of each event is recorded sequentially as dummy variable at each observed time point. Since malaria transmission intensity is time dependant, modelling MTIMBA mortality data using logistic regression models is appropriate. Mortality data is linked to predict EIR using a minimum distance approach. The analysis should consider EIR as a current exposure or as a cumulative exposure. Prior to mortality model fit, exploratory analysis using Kaplan-Meier survival curves, Log-rank and Wilcoxon tests need to be considered.

2.4.5 Model validation

Model validation is dependent on the type of models that have been used in estimating exposure effect. For MTIMBA data models fitted included, non-spatial, spatial, temporal and spatio-temporal(Amek et al., 2011; Amek., 2013; Kasasa, et al.,2013; Rumisha 2013) . Different methods that have been used in determining model's predictive ability including; Kullback-Leibler divergences, mean absolute error, chi-square, and credible interval plots (Schur et al., 2011). Model fit was carried out on a randomly selected sample (85%) of the data (training sample) and the remaining set was used for validation (test sample) .The best model was used in predict outcomes for the entire study area.

2.5 Example: spatio-temporal modelling of sporozoite rate data

For the MTIMBA data, zero-inflated and standard Bayesian logistic regression, Poisson/negative binomial models were fitted and tested by site. In this paper, we applied zero-inflated logistic regression models to analyze sporozoite rate data from Manhiça DSS,

Entomological data were obtained from 2918 georeferenced compounds in Manhiça DSS where light traps were set between October 2001 and September 2004. Fourty eight percent of the locations had mosquitoes for testing. A total of 1393 traps where *An. funestus* were identified, 89.4% had no infected mosquitoes. *An. gambiea* traps with no infected mosquitoes accounted for 94.2%. The DSS is located in the district of Manhiça (Maputo Province) in southern Mozambique. A full description of geographical and other characteristics of the area has been documented elsewhere (Aranda et al., 2005) .

2.5.1 Environmental data

Remote sensing data were downloaded from various sources at defined resolutions (Table 2.1). The climatic and environmental covariates were extracted at the locations where entomological data were available. For each location, temperature, rainfall and vegetation data were summarized by month for each year of the project.

Table 2. 1: Environment and climatic data source

Predictor	Spatial Resolution	Temporal Resolution	Source
Day land surface temperature (Day LST)	1 km²	8 days	MODIS
Night land surface temperature (Night LST)	1 km²	8 days	MODIS
Normalized difference vegetation index (NDVI)	250 m²	16 days	MODIS
Enhanced Vegetation Index (EVI)	250 m²	16 days	MODIS
Rainfall estimate (RFE)	8 km²	Dekadal	ADDS
Elevation/Altitude	1 km²	-	USGS
Nearest distance to water bodies (rivers and wetlands)	-	-	Local and Health Mapper

2.5.2 Model fit and implementation

All the data at each location were collapsed by month. Non-spatial analysis was conducted in STATA to assess the effect of elapsing time (lags) using Akaike's information criterion (AIC). For temperature three proxies were considered; land surface temperature day, night and average temperature. AIC was used to identify a suitable combination of climatic and environmental predictors for both sporozoite rate and density by vector species.

Bayesian geostatistical ZIB regression model was then fitted to sporozoite rate data. Location specific random effects were included in order to account for spatial heterogeneity. The covariance between any pair of locations was assumed to be an exponential function of distance between each pair of locations. A first-order autoregressive term was further added to the model in order to account for the temporal effect. Non spatio-temporal variation (nugget parameter) in the data was accounted for by an additional set of random effects which were considered as mutually independent and normally distributed with zero mean. Those random effects and a set

of covariates were modelled on a logit scale. Details of mathematical description for the model used are given in appendix 1

2.5.3 Results

A total of 18923 mosquitoes from 1445 compounds that were tested for *Plasmodium falciparum*, 16078 (85%) were *An. funestus* and rest (15%) were *An. gambiae*. The overall sporozoite rate accounted for 1.3% of the total mosquitoes trapped in 3 years. *Plasmodium falciparum* infections were detected in 1.4% of *An. funestus* and 1.1% in *An. gambiae*. There were more infected mosquitoes in the wet season (1.6%) than in the dry period (1.1%). Annual sporozoite rates from year one to three were 1.5%, 1.8% and 0.3% respectively. Figures 2.1 and 2.2 show monthly sporozoite rates by mosquito species. There were more infectious An. *funestus* between October and April, the wet and warm season. There was almost no infected An. *gambiae*, trapped in the warm months of November and December [Figure 2.2].

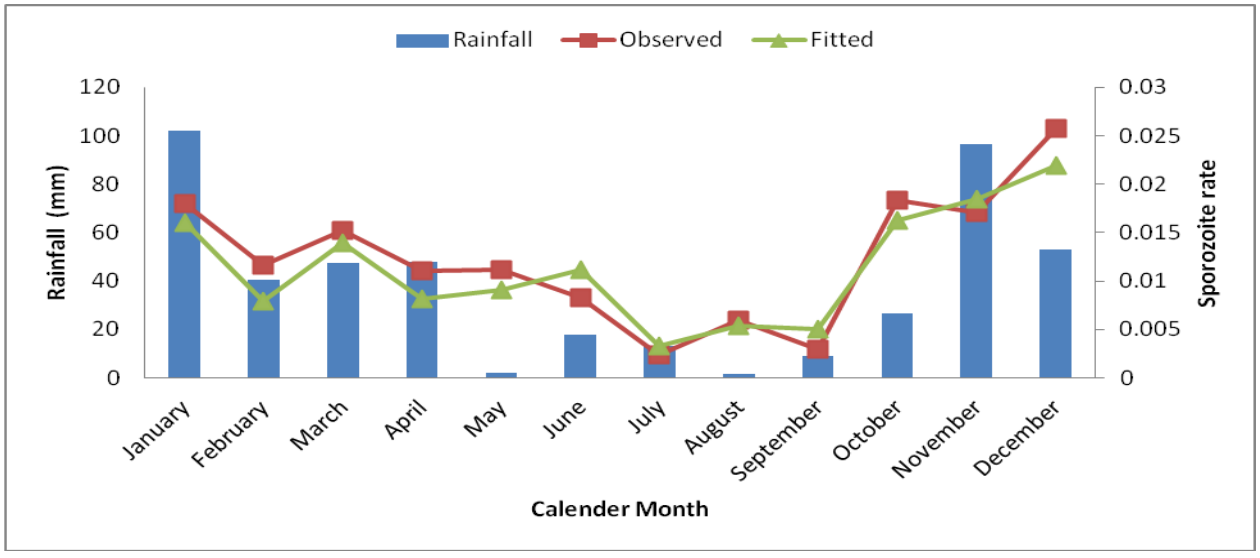


Figure 2. 1: Monthly rainfall, observed and fitted *An. funestus* sporozoite rates

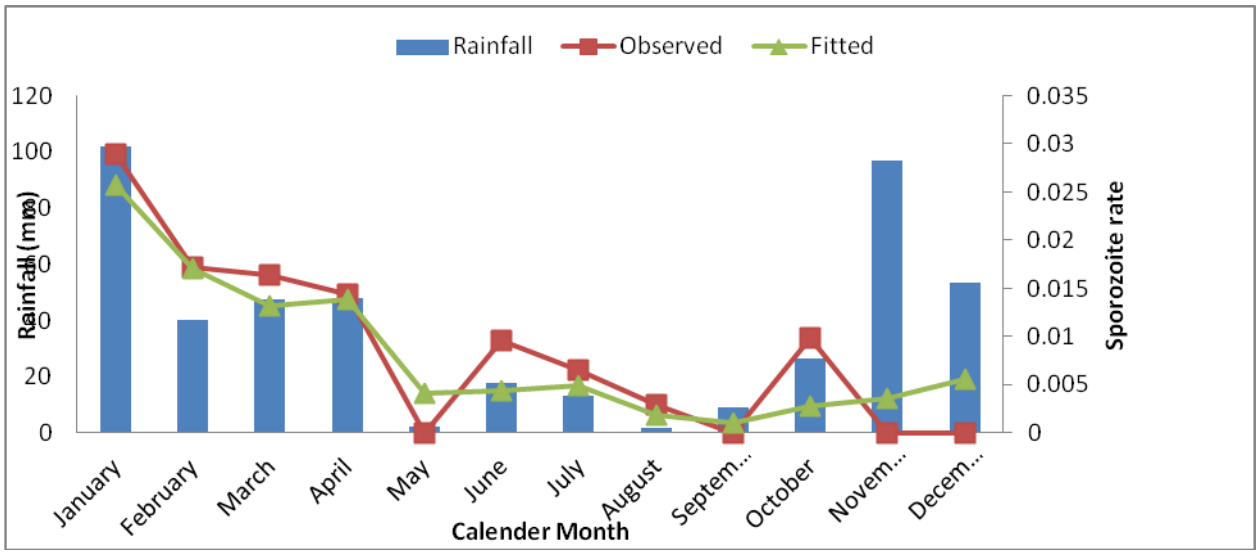


Figure 2. 2: Monthly rainfall, observed and fitted *An. gambiea* sporozoite rates

Analysis suggested that, the best combination of environment factors for An. *funestus* were; night LST of the survey month, day LST and NDVI of the previous month plus rainfall for two months before. *An. gambiae* sporozoite rate environmental predictors were current day LST; NDVI and rainfall of previous month plus average of the current and previous month night LST. Figures 2.3 shows validation results with the proportion of test locations with sporozoite rate that were predicted by three different models. Within a 95% credible interval the non-spatial, spatial and spatio-temporal models included correctly 23%, 45% and 51% test locations respectively. Apart from the 90% credible interval, the spatio-temporal model has been consistently predicting correctly more locations than both the spatial and non-spatial models. Based on the validation results, the spatio-temporal model was finally used to predict sporozoite rate by specie at un-sampled locations.

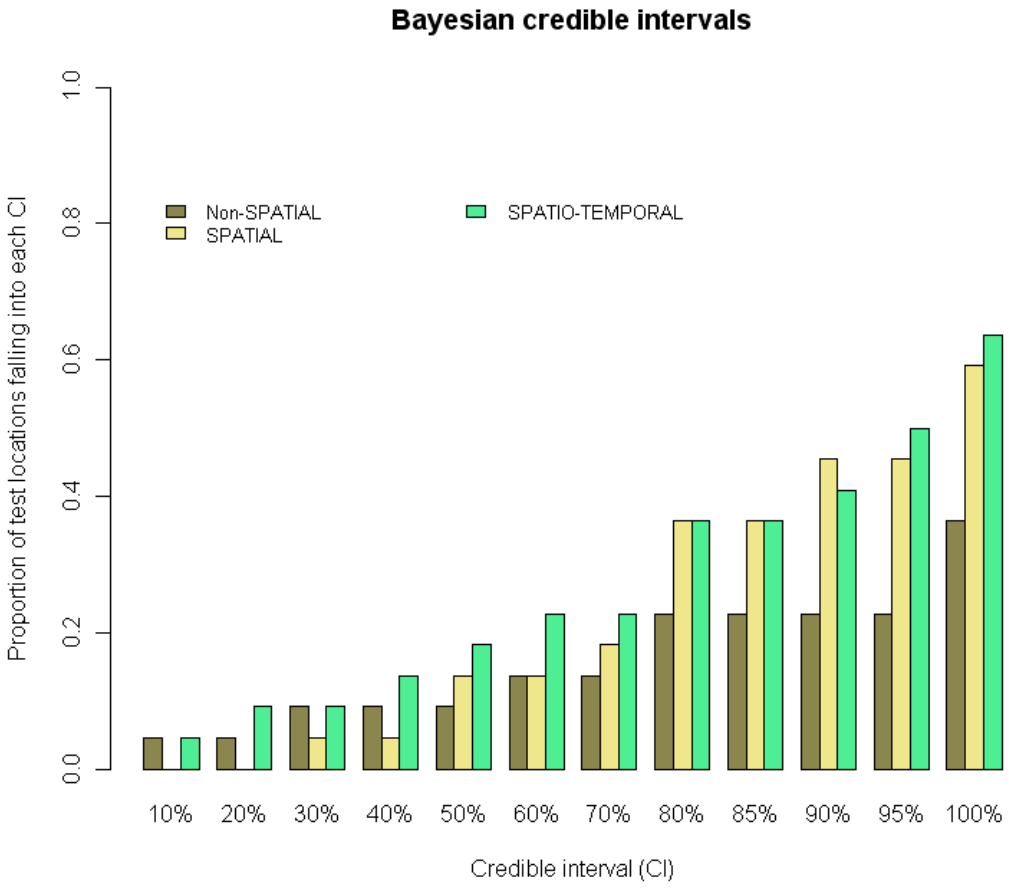


Figure 2. 3: Credible intervals of the posterior predictive distribution

Zero-inflated spatio- temporal model without an independent error term had the best fit for the two species [Figures 2.2 and 2.3] and posterior estimates taking into consideration the effect of climate and environmental factors are included in the Table 2.2.

Table 2. 2: Multivariate spatio-temporal analysis of sporozoite rate by mosquito species

Characteristics	*An. funestus*		*An. gambiae*	
	Co-efficients		Co-efficients	
	Median[a]	95% CI	Median[a]	95% CI
Intercept	-3.50	(-7.88, 1.45)	-2.10	(-7.68, 4.15)
Altitude	0.003	(-0.01, 0.02)	0.01	(-0.02, 0.04)
Distance to water bodies	-1. 23	(-0.46, 0.17)	-0.04	(-0.69, 0.46)
NDVI	-0.18	(-3.37, 2.86)	0.21	(-4.46, 5.06)
Rainfall	0.01	(-0.01, 0.04)	-0.004	(-0.04, 0.03)
Season(Wet)	0.12	(-1.00, 1.47)	1.91	(-0.35, 4.59)
Day temperature	-0.07	(-0.26, 0.11)	0.23	(-0.005, 0.66)
Night temperature	0.03	(-0.24, 0.27)	-0.50	(-1.07, -0.01)
Variances				
Spatial (σ_{ϕ}^2)	0.03	(0.16, 0.69)	0.43	(0.16, 1.34)
Temporal (σ_{ε}^2)	0.47	(0.18, 1.52)	0.49	(0.17, 2.25)
Range (in km)	1.57	(0.33, 7.84)	0.67	(0.34, 5.60)

[a] :Median of the posterior distribution using ZIB model

Multivariate analysis shows that altitude, rainfall, season, night temperature were positively associated with *An. funestus* sporozoite rate. A negative association with the same specie was observed between closest distances to water bodies, vegetation (NDVI) and day temperature. All the covariates were not significantly associated with *An. funestus* sporozoite rate. The estimated temporal variance parameter (σ_{ε}^2 =0.5, 95% CI: 0.18, 1.52) is larger than spatial variance (σ_{ϕ}^2 =0.03, 95% CI: 0.16, 0.69). The minimum distance at which the correlation becomes negligible is 1.6Km (95% CI: 0.3km, 7.8km).

An. gambiae, sporozoite rates were positive associated with altitude, vegetation index (NDVI), wet season and day temperature. Similarly, distance to water bodies, rainfall and night temperate were negatively associated with *An. gambiae* sporozoite rate. Apart from night temperate, all other predictors were not significantly related to sporozoite rate. From *An. gambiae* data, the estimated temporal variance parameter (σ_{ε}^2 =0.49, 95% CI: 0.17, 2.25) is slightly larger than

spatial variance (σ_{ϕ}^{2} =0.43, 95% CI: 0.16, 1.34). The minimum distance at which the correlation becomes negligible is 0.7Km (95% CI: 0.3km, 5.6km), indicating a weak correlation in sporozoite data.

Combined predicted sporozoite rates for the two species covering the entire DSS by month are shown in Figure 2.4 The figure shows a higher prediction of sporozoite rate between October and February, the wet and warm period. The predicted errors with specific locations sampled in a month are shown in Figure 2.5. The period May to October shows higher prediction errors because this is a season when even observed sporozoite rates are close to zero (Figures 2.1 and 2.2).

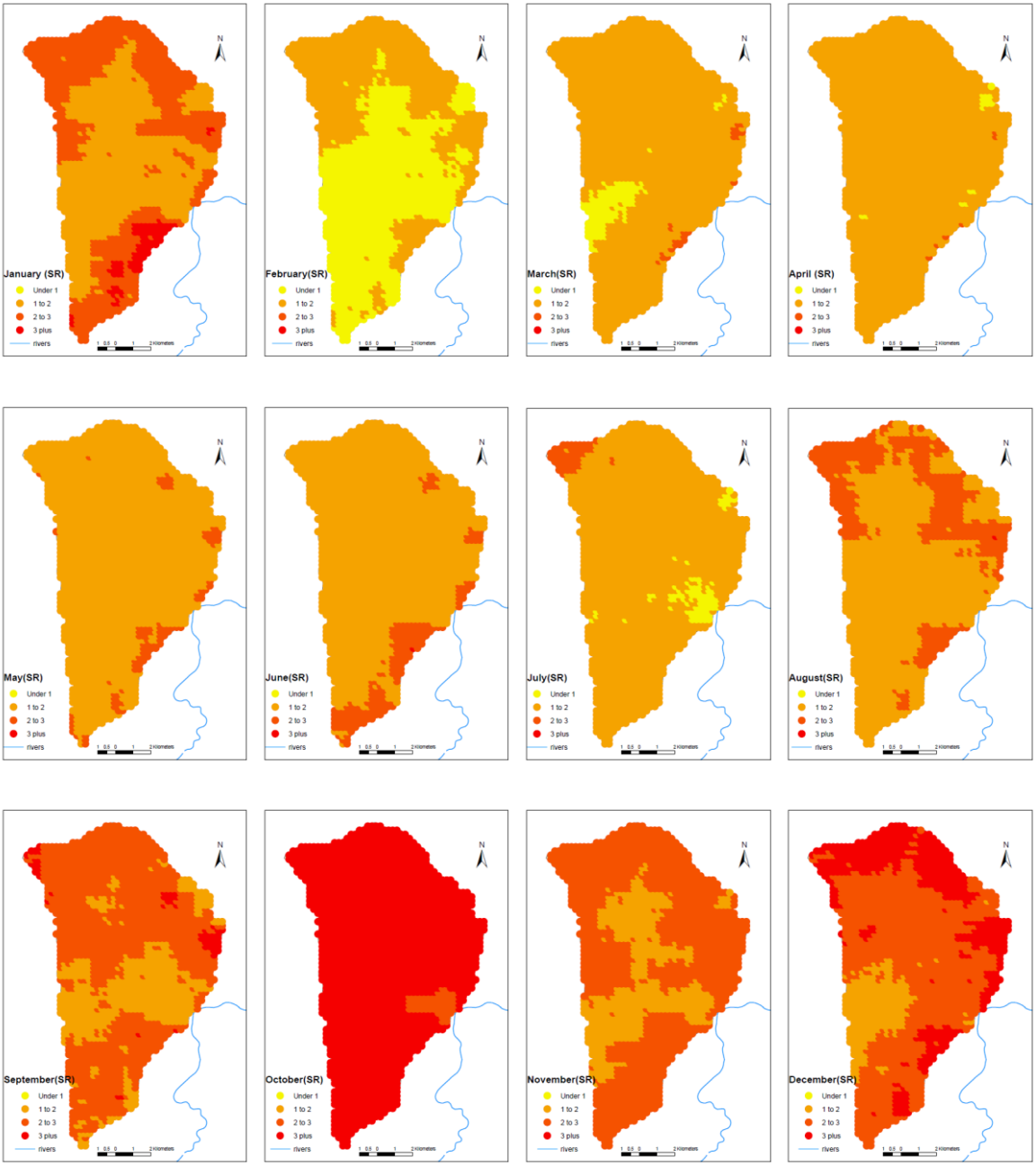


Figure 2. 4: Predicted sporozoite rates from geostatistics models

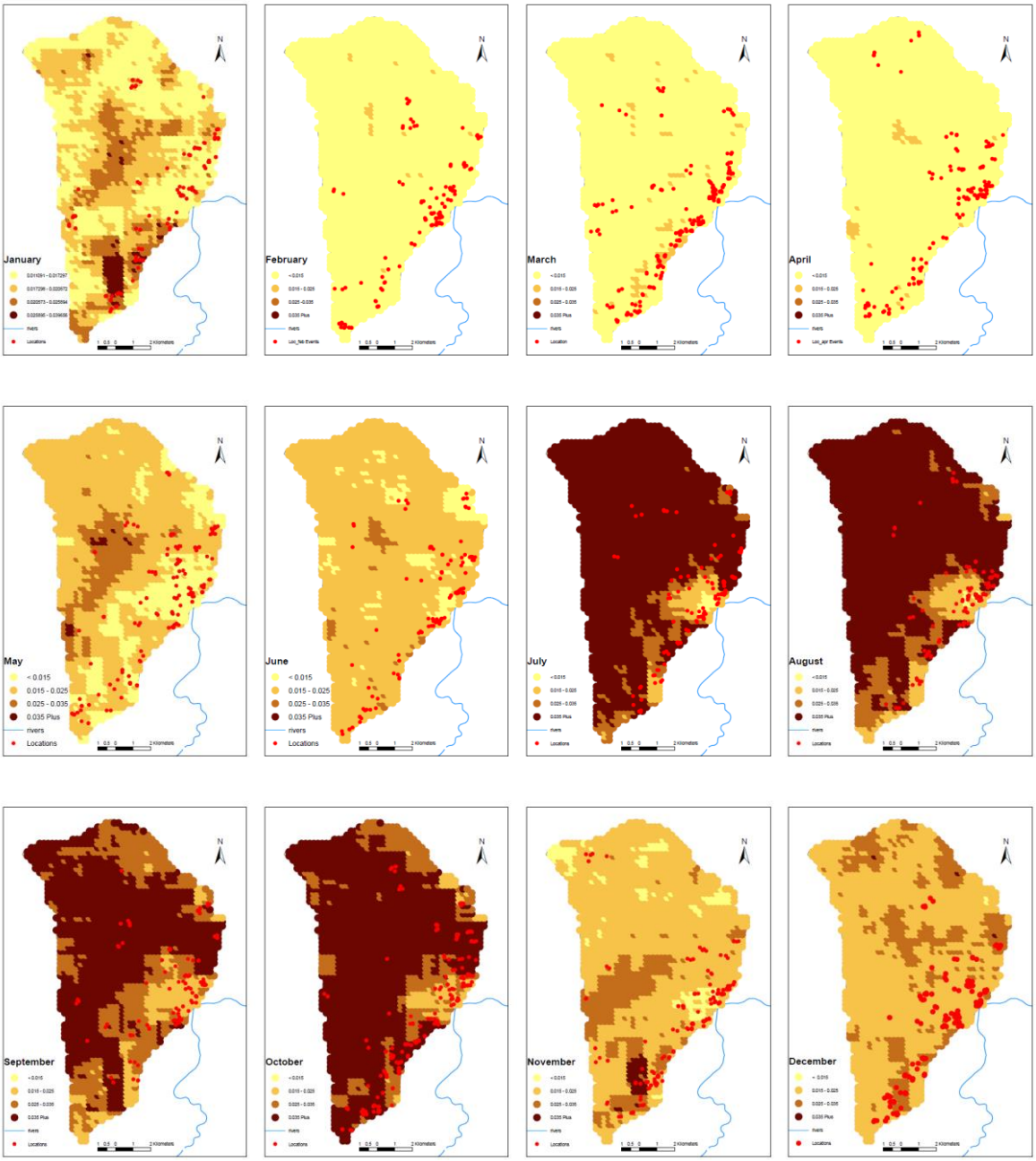


Figure 2. 5: Monthly prediction errors

2.5.4 Discussion

In this paper, we give an overview of the MTIMBA project and analyzed space time correlated sporozoite rate data from Manhiça DSS. To our knowledge, INDEPTH is the first to assemble the largest entomological database in Africa. Although generating entomological data is costly,

EIR is the recommended measure of transmission intensity in malaria endemic countries. The network initiated the project in early 2000 in order to further analyze the relationship between malaria transmission intensity and mortality (Gemperli et al., 2004; Smith et al., 2001). The project sites completed data collection process in 2004. Merging all these data into a single database was done in 2005. While analyzing MTIMBA data, the EIR was estimated from zero inflated Bayesian geostatistical models fitted separately for both sporozoite rate and mosquito density data. This database can also be used to assess malaria epidemiology especially disease heterogeneity within and between sites.

Analysis plans for the MTIMBA data included a capacity building component targeting people from malaria endemic area. Five students from four DSS (Ifakara, Iganga , Kisumu and Nouna) (Adazu et al., 2005; Hammer et al., 2006.; Rutebemberwa et al., 2009; Somi et al., 2007) and one national research organization were admitted for doctoral studies at Swiss Tropical and Public Health Institute (Swiss TPH), Switzerland. Through course work, seminars and mentoring from experts at the Swiss TPH and MTIMBA sites, these students have been able to acquire knowledge in malaria epidemiology especially in the areas of Bayesian geostatistical modelling and disease mapping. It is expected that the statistical and writing skill acquired will then be transferred back to SSA sites. Application of spatio-temporal analysis techniques will be useful in timely identification of disease hot spots for optimal allocation of scarce resources. Training is likely to promote further research within the INDEPTH network and also between north and south.

Like any other longitudinal study, MTIMBA project faced challenges ranging from logistical support, manpower to technical expertise especially in the area of statistics. Due to climatic changes, sites were enabling to access some of the pre-selected compounds leading to increased

number of zero mosquito catches in density data. Although the project developed a protocol, sites had to use resources and means available to them in order to generate the expected data. For instance, Navrongo HDSS use two laboratories to test mosquitoes for *Plasmodium falaciparium* circumsporozoite protein. These laboratories seemed to have used different cut-off points leading to large variation in sporozoite rates between the first and proceeding years (Kasasa Simon, et al., 2013). Although site specific VA questionnaires were recommended in the protocol, cause specific mortality data for all sites were incomplete. In addition all compounds in the participating sites were not geo-referenced. This reduced the number of locations to include in the spatial analysis for both entomological and mortality models.

Tools for collecting complementary data on interventions, social economic status were not standardized. Some sites provided data for mosquito nets ownership while others on use. Some sites did not collect these data at all.

Although MTIMBA projected registered certain gaps in the data, outputs generated from within and between site analyses will form benchmarks in monitoring and evaluation of malaria control intervention.

Results from our spatio-temporal models predicted the highest number of locations between 10% and 95% confidence interval for the sporozoite data (Figure 2.1). From multivariate analysis, distance to water bodies was the only single factor that was negatively associated to the two mosquito species. These distances were computed from locally generated maps of flood areas and swamps which are the bleeding sites for mosquitoes. Such places are close to compounds which give mosquitoes easy access to a blood meal. The shortest distance at which the spatial correlation was below 5% was low (between 3 to 8 km) for the two species. This shows a faster decay of the correlation with distance for sporozoite rate data. The weak correlation in data

seems to suggest that sporozoite rate depends largely on local conditions rather than environmental factors over the large area. The weak correlation in sporozoite rate data was also observed in northern Ghana (Kasasa et al., 2013).

Wet seasons in Manhiça had a positive relationship with sporozoite rate. This is also evident from figures 2.1 and 2.2. November to April are the warm and wet months for the area (Aranda et al., 2005). This period creates a favourable environment for mosquito breeding and survival.

Analyses of the Rufiji and Kisumu MTIMBA-DSS data have been completed. Work on the data from the remaining sites (Manhiça, Navrongo, Ifakara and Nouna) is still on-going. Although Rufiji experiences higher transmission intensity than Kisumu, results indicate the presence of temporal and season variation in both sites. A strong association between all-cause mortality and malaria transmission was observed among children (1-4 years) and school age children in Kisumu and Rufiji respectively(Amek, 2013; Rumisha, 2013). A negative association between mortality and transmission was observed among old people in the two sites. Completing site-specific analysis will help in generating a firm conclusion that will direct malaria control interventions.

2.6 Appendix 1

Spatio-temporal modelling of sporozoite rate

Let N_{it} be the number of mosquitoes tested by vector species and Y_{it} be the number of positive mosquitoes at location s_i, $i = 1,....,n$ and time t. We assumed that Y_{it} follows a binomial distribution; that is $Y_{it} \sim Bin\left(N_{it,} p_{it}\right)$ with parameter p_{it}, the sporozoite rate. To account for spatial, temporal and random variation in the data, we introduced location, monthly and independent random effects ϕ_i , ε_t and e_i respectively. These were modelled together with covariates X_{it} via the logistic regression $\log it(P_{it}) = X_{it}^{T}\beta + \phi_i + \varepsilon_t + e_i$, where $\beta = (\beta_1, \beta_2, ..., \beta_k)^T$ is a vector of regression coefficients. The spatial random effects are assumed to originate from a Gaussian spatial process with zero mean and variance-covariance matrix Σ , where $\phi_i = (\phi_1, \phi_2, ..., \phi_n)^T \sim N(0, \Sigma)$. The covariance between any particular pair of locations was assumed to be a function of distance between the locations, that is $\Sigma_{ij} = \sigma_\phi^{\,2} \exp\left(-d_{ij}\ \rho\right)$ where d_{ij} is the Euclidean distance between locations s_i and s_j, $\sigma_\phi^{\,2}$ is the spatial variance and ρ is the correlation decay parameter with the range defined as $3/\rho$ (Ecker and Gelfand, 1997). Temporal random effects were modelled by first order autoregressive process [AR (1)] with variance $\sigma_\varepsilon^{\,2}$ which allows correlation between consecutive time periods(J L Hay and Pettitt, 2001). Non spatio-temporal random effects e_i were assumed to follow a normal distribution with zero mean, $e_i \sim N(0, \sigma_e^2)$.

Model fit

Following Bayesian model specification, various priors for model distribution parameters were adopted. Prior distributions included; normal with zero mean and large variance for regression coefficients $[\beta \sim N(0, 10^2)]$, inverse gamma for spatial and independent error

variances $[(\sigma_{\phi}^2, \sigma_e^2);\ \sigma^2 \sim IG(2.01, 1.01)]$, uniform for decay parameter $[\rho \sim U(a\ , b\ )]$ and beta for the mixing proportion $[\pi \sim Be(1,1)]$.

Markov Chain Monte Carol (MCMC) simulation algorithm was used in estimating the model parameters. We used a single chain sampler of 180000 iterations with an initial burn-in of 10000. OpenBUGS software was used for parameter estimation. A FORTRAN program written by the authors was used for Bayesian Kriging in order to predict sporozoite rate by species at unsampled locations (Diggle et al., 1998) . Convergence was assessed after running long chains and keep on monitoring trace and density plots.

Model validation

Three types of models, namely non-spatial, spatial and spatio-temporal were fitted. Using balanced sampling, 85% of the data was selected as a training set for model fit; while the rest (15%) were used for validation in order determine the models predictive performance. Observed sporozoite rate at test locations and predictions were compared for accuracy using 95% Bayesian credible interval approach (Schur et al., 2011). A model with highest percentage of locations within the credible interval was assumed to have the best predictive ability.

Chapter 3: Spatio-temporal malaria transmission patterns in Navrongo Demographic surveillance site, Northern Ghana

Authors: Simon Kasasa[1,2,3], Victor Asoala[4], Laura Gosoniu[1,2], Francis Anto[5], Martin Adjuik[6], Cletus Tindana[4], Thomas Smith[1,2], Seth Owusu-Agyei[7], Penelope Vounatsou[1,2 *]

[1]Swiss Tropical and Public Health Institute, Socinstrasse 57, P.O. Box 4002 Basel, Switzerland,
[2]University of Basel, Switzerland
[3]School of Public Health, Makerere University college of Health Sciences, Uganda
[4]Navrongo Health Research Centre, Ghana
[5]School of Public Health, University of Ghana, Legon, Ghana
[6]INDEPTH Network Secretariat, Accra, Ghana
[7]Kintampo Health Research centre, Ghana Health Services, Ministry of Health

*Corresponding Author

This paper has been published in Malaria Journal 2013, **12**:63

Abstract

Background: The relationship between entomological measures of malaria transmission intensity and mortality remains uncertain. This is partly because transmission is heterogeneous even within small geographical areas. Studying this relationship requires high resolution, spatially structured, longitudinal entomological data. Geostatistical models that have been used to analyse the spatio-temporal heterogeneity have not considered the uncertainty in both sporozoite rate (SR) and mosquito density data. This study analysed data from Kassena-Nankana districts in northern Ghana to obtain small area estimates of malaria transmission rates allowing for this uncertainty

Methods: Independent Bayesian geostatistical models for sporozoite rate and mosquito density were fitted to produce explicit EIR estimates for small areas and short time periods, controlling for environmental factors.

Results: Mosquitoes were trapped from 2,803 unique locations for three years using mainly CDC light traps. *Anopheles gambiae* constituted 52%, the rest were *Anopheles funestus.* Mean biting rates for *An. funestus* and *An. gambiae* were 32 and 33 respectively. Most bites occurred in September, the wettest month. The sporozoite rates were higher in the dry periods of the last two years compared with the wet period. The annual EIR varied from 1,132 to 157 infective bites. Monthly EIR varied between zero and 388 infective bites. Spatial correlation for SR was lower than that of mosquito densities.

Conclusion: This study confirms the presence of spatio-temporal heterogeneity in malaria transmission within a small geographical area. Spatial variance was stronger than temporal especially in the SR. The estimated EIR will be used in mortality analysis for the area.

Keywords: Entomological inoculation rate, Spatio-temporal, Zero-inflated, Malaria, Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project

3.1 Introduction

Malaria continues to be endemic in most sub-Saharan countries, particularly in Ghana where this study was carried out(Carneiro et al., 2010; Clerk et al., 2009; Oduro et al., 2007; "WHO | World Malaria Report 2009,") Malaria in Ghana is transmitted by two main vectors: *Anopheles gambiae* and *Anopheles funestus*, whose peak activities occur at the end of the wet season. Changes in climate, land use and environmental factors profoundly influence the vector, and hence the parasite and transmission patterns. Malaria transmission intensity is measured using clinical (spleen rate), parasitological (parasite infection rate), entomological (entomological inoculation rate [EIR]) or serological markers("A research agenda for malaria eradication," 2011; Drakeley et al., 2005). The most direct measurement of transmission intensity is EIR, the number of infective bites per person per unit time. It is calculated as a product of the proportion of mosquitoes with sporozoite in their salivary glands (sporozoite rate) and numbers of vectors biting an average human in unit time (the human biting rate)(Beier et al., 1999).

Malaria transmission in sub-Saharan Africa is heterogeneous, varying between climatic seasons, ecological zones and even among areas in close proximity(Carter et al., 2000; Charlwood et al., 1995; de Souza et al., 2010; Drakeley et al., 2003; Kelly-Hope and McKenzie, 2009; Mabaso et al., 2007; Okello et al., 2006; Shililu et al., 2003). In Ghana, malaria transmission has shown a clear variation over time, season and space(Abonuusum et al., 2010; Appawu et al., 2004; Dery et al., 2010). The relationship between malaria transmission and mortality is still unclear(Gemperli et al., 2004; Smith et al., 2001). To clarify the relationship between malaria transmission and mortality, the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project was established in 10 INDEPTH network sites between 2001 and 2004(Abdullah et al., 2007; Amek et al., 2012). Entomological data were collected every two

weeks over a large number of compounds within each site for a period of three years. Each site used a slightly different sampling strategy for mosquitoes depending on available resources and local settlement patterns, aiming to obtain an unbiased estimate of the numbers of biting mosquitoes. These data are spatially correlated because neighbouring compounds share common exposures such as interventions, land use, climate and environmental factors. The longitudinal nature of the data also introduces a temporal correlation.

Rumisha and Amek(Amek, 2013; Rumisha, 2013) developed geostatistical temporal models to obtain EIR exposure surfaces for the Rufiji and Kisumu MTIMBA-health and demographic surveillance (HDSS) sites, respectively. Subsequent analyses linking mortality to EIR exposure indicated a positive linear relationship between mortality and malaria transmission intensity among the under-fives and a negative association for individuals aged 60 years and above. Although malaria is common in sites, their endemicity, spatio-temporal patterns and mosquito composition are completely different. Malaria transmission in Rufuji is driven by both *An. funestus* and *An. gambiae*, while the later is dominant in Kisumu throughout the year. Kisumu experiences two transmission peaks in a year and Rufiji has only one. This is partly due to ecological differences between the two sites. In relation to breeding sites, *An. funestus* prefer clear, permanent fresh waters while *An. gambiae* larvae are found mostly in temporal and shallow water bodies. Estimating site-specific heterogeneity in malaria transmission will help clarify how variation in transmission influences the malaria-related mortality.

This study reports spatially and temporally explicit estimates of EIR at high resolution, obtained by analysing the MTIMBA data collected from Kassena-Nankana district in northern Ghana where the Navrongo health and demographic surveillance system (NHDSS) is located. The EIR

was estimated from Bayesian geostatistical models, fitted separately for sporozoite rate (SR) (assumed to be binomially distributed) and mosquito density data (negative binomially distributed). Model-based predictions at unobserved locations generated spatially explicit and season-specific estimates of EIR for the entire area. These estimates will subsequently be used in addressing the MTIMBA project's main objective of estimating the relationship of mortality with malaria transmission.

3.2 Methods

3.2.1 Description of the Study area

The NHDSS is located in the administrative district of Kassena-Nankana (between latitude 10^0 $30^{'}$ and 11^0 $00^{'}$ North and longitude 1^0 $00^{'}$ and 1^0 $30^{'}$ West), in northern Ghana, bordering Burkina Faso. Its altitude stretches up to 400 m above sea level. The district covers an area of 1,675 sq km and lies within the Guinea savannah belt. Approximately 140,000 people reside in the district and the majority is subsistence farmers. There are two distinct seasons; the wet, between April and October and a dry period that covers remaining months of the year. The region receives approximately 850 mm of precipitation per year with monthly temperatures ranging between 18 °C and 45 °C. The HDSS routinely collects demographic data using "a compound" as a unit of observation. Malaria is endemic in the area and *Plasmodium falciparum* is transmitted by both *An. gambiae* and *An. funestus*. *Anopheles gambiae* s.s. has previously been reported as a dominant sibling species of the *An. gambiae* complex. The *An. gambiae* M form is predominant in the northern parts of Ghana where NHDSS is located(Charlwood et al., 1995; de Souza et al., 2010). The canals from Tono dam and irrigated lands serve as breeding sites for *An. gambiae* throughout the year, while the rice fields support *An. funestus* breeding especially

during the periods when the vegetation is flooded. The small dams that are used in the dry seasons favour mosquito growth in these areas. Malaria transmission in the district occurs throughout the year. Between 2001 and 2002, the recorded mean EIR for the district was as high as 418 infective bites per person per year (ib/p/y)(Appawu et al., 2004). Further characteristics of the district and the HDSS have been described elsewhere(Appawu et al., 2004; Oduro et al., 2007).

3.2.2 Data types and sources

i) Entomological data

Mosquitoes were collected from randomly selected compounds using both light traps and human landing methods following the MTIMBA protocol. Compounds were randomly selected at the beginning of the study using the HDSS database and were allocated to trapping weeks. Sampled compounds were between 100–500 meters apart and were balanced in terms of numbers for the two major zones namely: irrigated and non-irrigated areas. Only one trap was set per compound per night. Light trap catches were performed overnight (from 18:00 GMT to 06:00 GMT). No study team member visited the compound at night until it was time to remove traps the next morning. Such visits were perceived by community members as intrusion. Traps were hung about 1.5 m above the floor next to the bed of an "indexed" person. Heads and thoraces of light-trapped *Anopheles* were tested for *P. falciparum* circumsporozoite protein using enzyme linked immunosorbent assay (ELISA) (Wirtz et al., 1987).

The entomological inoculation rate was therefore computed as a product of human biting rate and the proportion of infectious mosquitoes (sporozoite rate). Human biting rate was estimated as a geometric mean of *Anopheles* mosquitoes caught per light trap set(Lines et al., 1991).

Mosquitoes were trapped in 56% of the 2,803 uniquely georeferenced compounds within the site. Infectious mosquitoes were only found in 28% of these locations.

ii) Environmental data

Environmental and climatic predictors were obtained from various remote sensing sources. Day and night land surface temperature (LST) at 1 x 1 km and both normalized difference vegetation index (NDVI) plus enhanced vegetation index (EVI) at 250 x 250 m were downloaded from Moderate Resolution Imaging Spectro-radiometer (MODIS). LST and vegetation data were extracted at eight-day and 16-day temporal resolutions respectively. Rainfall estimates (RFE) at 8 x 8 km were obtained at 10-day intervals from the African Data Dissemination Service (ADDS). Altitude at 1 x 1 km was obtained from US Geological Survey (USGS) data centre. Distance to water bodies (based on local rivers and wetlands) was downloaded from HealthMapper version 4.2 databases. The shortest Euclidean distance from water bodies to compounds was calculated using ArcGIS version 9.1 software. The climatic and environmental variables were processed at the locations where entomological data were available. For each location, temperature, rainfall and vegetation data were summarized by month for each year of the project.

3.2.3 Data analysis

Non-spatial logistic and negative binomial regression models were used to analyse sporozoite and density data respectively. Zero-inflated models were fitted to account for the large number of locations with either no mosquitoes (44%) or no infectious mosquitoes (72%). The Akaike's information criterion (AIC) in STATA was used to assess the length of the elapsing time (lags)

between climatic suitability and malaria transmission. In particular, five summary estimates were computed for each of the environment factors based on mosquito collection month in a year: i) current month of collection, ii) previous month, iii) previous two months, iv) average of the current and previous month, and v) average of the current and previous two months. Three temperature proxies were considered: land surface day, night and average temperature. Seasonality was taken into account by either a binary variable (wet/dry) or trigonometric functions with: (i) one cycle indicating a single transmission season, or (ii) two cycles corresponding to two transmission seasons per year. AIC was used to identify a suitable combination of climatic and environmental predictors for both SR and density by vector species.

Bayesian geostatistical formulations of the above models were fitted to take into account spatio-temporal correlation. In each model, compound-specific random effects were included. They were assumed to be latent observations from a multivariate Gaussian spatial process with a zero mean. The covariance of the process included the spatial variance and an exponential correlation function of distance between any pair of compound locations. First-order autoregressive terms were included to model temporal correlation. Any remaining non-spatial variation (nugget parameter) was considered by an additional set of location random effects, assumed to be mutually independent and normally distributed with zero mean. All the corresponding random and the covariates effects were modelled either on a logit or log scale depending on the model; logistic regression for sporozoite and negative binomial regression for the mosquito density data, respectively. Bayesian kriging was applied to predict SR and mosquito density over a grid of 31,308 pixels with 250 x 250 m spatial resolution. The analysis was carried out for each mosquito species (i.e. *An. funestus* and *An. gambiae*) separately. Mosquito densities were converted to man-biting rates after adjusting for a factor. The indices were multiplied at each

location to generate spatially explicit surfaces of EIR for each species. Maps for the total EIR were generated using ArcGIS software. Details of mathematical description for all models used are given in appendix 2.

3.2.4 Model Validation

Models were fitted on 85% of the locations (training sample) and they were validated on the remaining 15% of locations (test sample). In particular, the model's predictive ability was assessed by estimating the proportion of test locations correctly predicted within Bayesian credible intervals of probability coverage varying from 1 to 100% (Gosoniu et al., 2006). The model with the highest number of correctly predicted locations consistently over the intervals was considered as the one with the best predictive performance.

3.3 Results

3.3.1 Description of density data

The mean biting rates per person and night for *An. funestus* were 34 in the first year, 32 in the second and 19 in the third. Similarly, *An. gambiae* mean bites were 33 in the first year, followed by 26 and 15 bites in the second and final year respectively. For the entire research period, mean biting rates per month varied with seasonal changes. For both species, most bites were observed during the wet season (July to November). Highest bites occurred in the month of September for all the three years. During the dry period of January to April, fewer monthly bites were recorded. Mosquitoes in the area became more abundant after the first three months of the rainy season [Figure3.1].

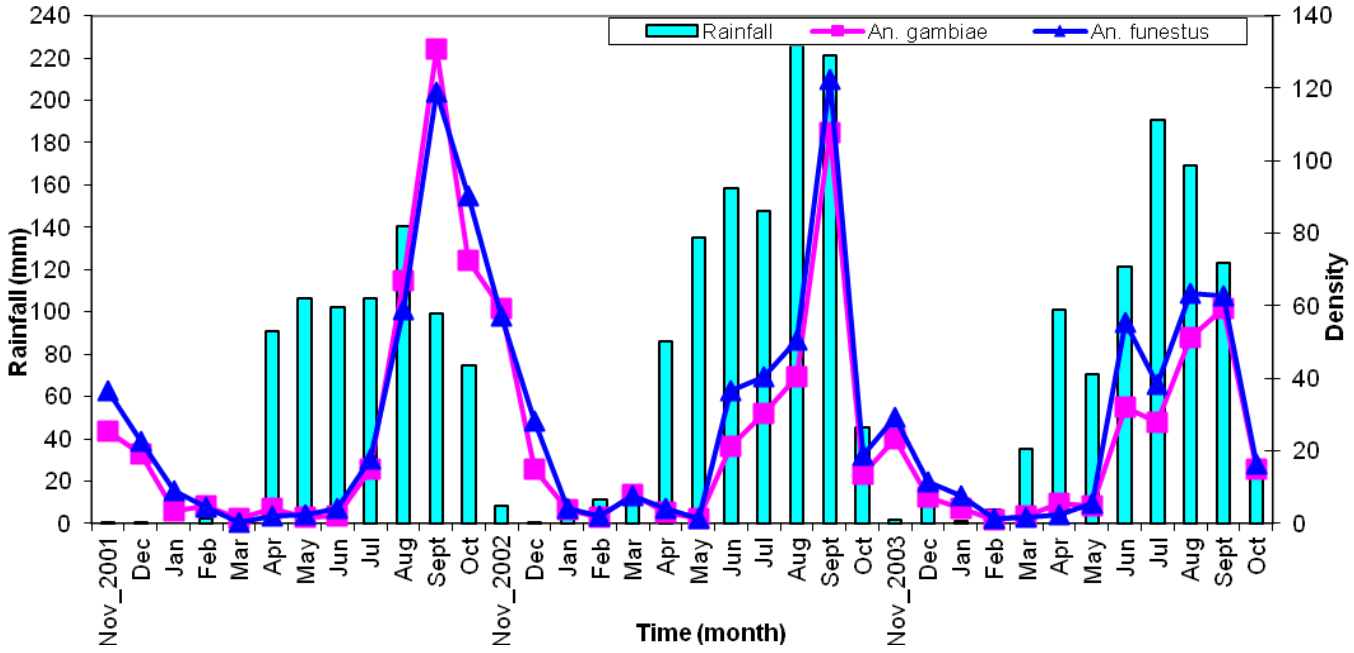


Figure 3. 1: Monthly rainfall and observed mosquito density

3.3.2 Description of sporozoite rate data

A total of 109,647 malaria mosquitoes from 1,565 compounds were tested for sporozoites; 56,887 (52%) were *An. funestus* and the rest were *An. gambiae*. The overall SR was 2.5%. *Plasmodium falciparum* infections were detected in 2.4% of *An. funestus* and 2.7% in *An. gambiae*. The proportion of infectious An. *funestus* was almost equal to that of *An. gambiae* in both the first (4.8% and 4.7%) and third (1.2% and 1.4%) years. The lowest SR of 0.8% was observed in the second year from *An. funestus* mosquito species. The data showed an overall SR of 1.8% and 2.7% in dry and wet season respectively. However, during the second year the dry period SR was more than double that of wet season (1.6% compared with 0.7%). The proportion of infectious *An. gambiae* (2.1%) was higher than that of *An. funestus* (1.5%) in the dry season.

The fraction of infected *An. gambiae* mosquitoes was higher in dry season than wet for the second (2.0%) and third (1.5%) year. The monthly SR for both species follows a similar pattern for all the three years [Figure 3.2].

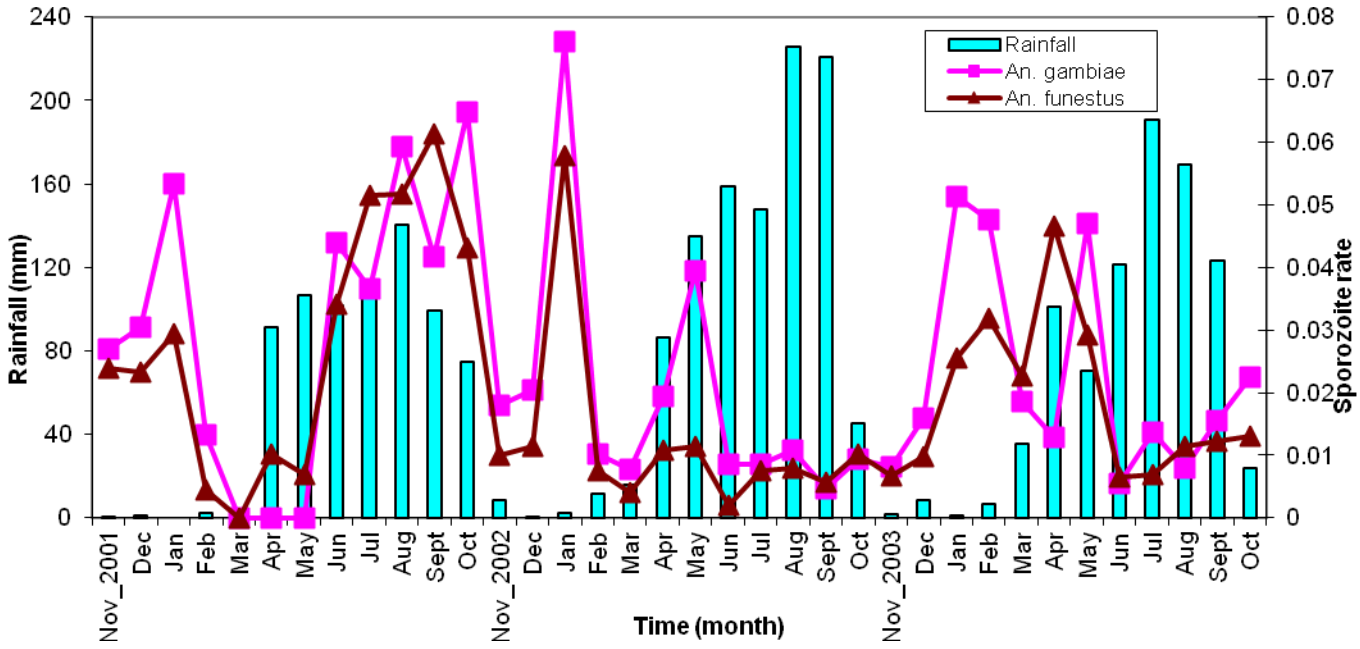


Figure 3. 2: Monthly rainfall and observed sporozoite rate by mosquito species

3.3.3 Description of EIR data

The crude annual EIR estimates, based on entomological data from first to third year were 1132, 193 and 157 ib/p/y respectively [Table3.1].

Table 3. 1: Observed entomological inoculation rate

	EIR per person per year		
Year	*An. funestus*	*An. gambiae*	Combined species
1	575	557	1132
2	90	103	193
3	79	78	157

The crude annual EIR estimates, based on entomological data from first to third year were 1132, 193 and 157 ib/p/y respectively (Table 1). The highest EIR was observed in the month of September of each year and varied from 388 in the first year to 37 and 51 infective bites per month in the second and third year respectively. For all the three years, lowest monthly infective bites were observed either in February or March [Figure 3.3].

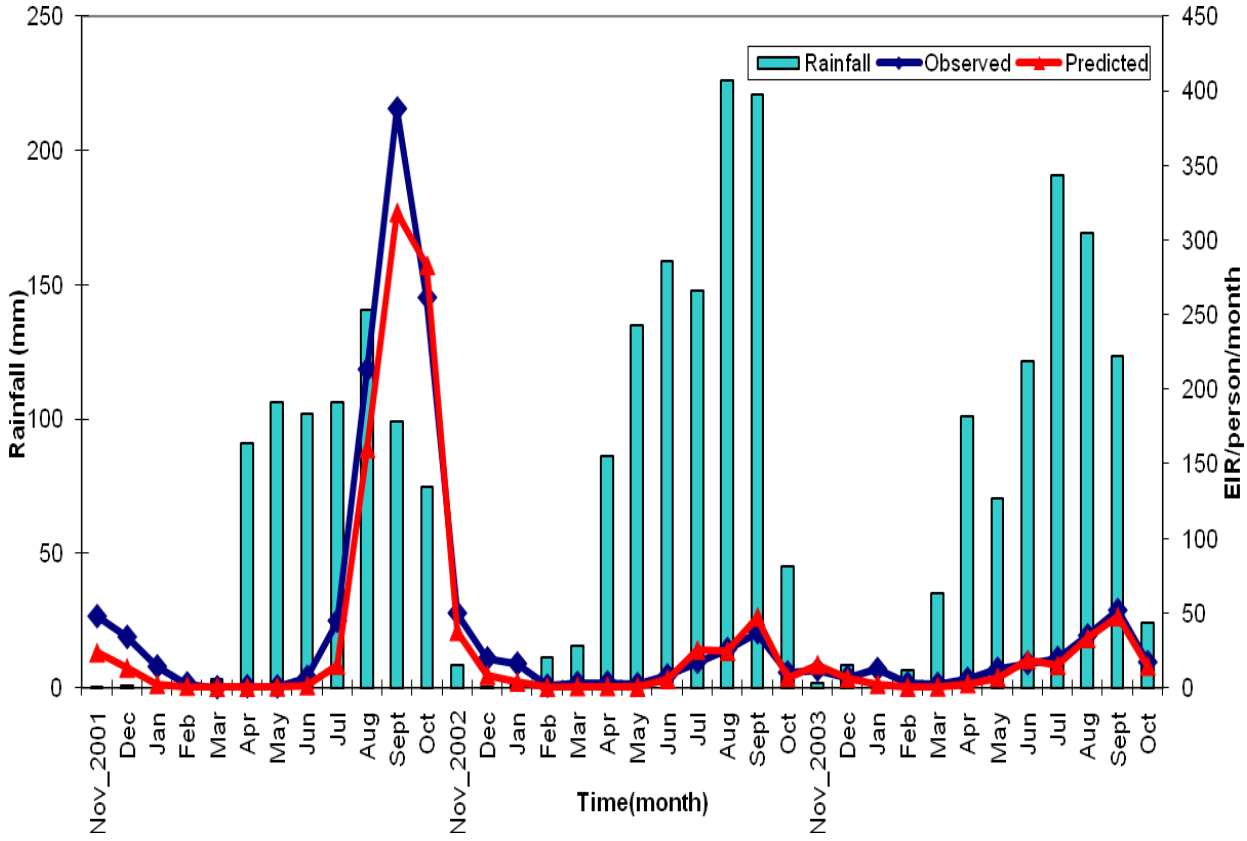


Figure 3. 3: Observed and predicted EIR

3.3.4 Model-based results: Mosquito density data

Lag time analysis showed that mosquito density for both species was related to current NDVI, total rainfall, average day and average night temperatures over the two months prior to the survey. Parameter estimates from geostatistical, zero-inflated, negative binomial models are summarized in Table 2. For *An. funestus*, distance to water bodies, NDVI, season, day

temperature and second year of data collection were related to density. An increase in vegetation cover was highly associated with an increase in biting rates. Compounds that are close to water bodies were associated with higher number of mosquito bites. Wet seasons and increase in day land surface temperatures were negatively associated with mosquito density. Spatial variation ($\sigma_{\phi}^{(D)2}$ =0.9, (95% CI: 0.5, 1.6)) was almost similar to the temporal one ($\sigma_{\varepsilon}^{(D)2}$ =0.82, (95% CI: 0.5, 1.5)).

Table 3. 2: Multivariate spatio-temporal analysis for mosquito density by species

Parameters	*An. funestus*		*An. gambiae*	
	Co-efficients		Co-efficients	
	Median	95% CI	Median	95% CI
Intercept	2.73	(2.16, 3.31)	1.86	(1.05, 3.67)
Altitude	-0.01	(-0.02, 0.00)	-0.01	(-0.02, 0.00)
Distance to water bodies	-0.12	(-0.22, -0.02)	-0.18	(-0.27, -0.07)
NDVI	2.27	(1.44, 2.83)	1.51	(1.17, 2.39)
Rainfall	0.002	(-0.003, 0.01)	0.0002	(-0.01, 0.01)
Season(Wet)	-0.23	(-0.61, -0.003)	-0.26	(-1.13, 0.33)
Day temperature	-0.04	(-0.09, -0.004)	-0.08	(-0.13, -0.04)
Night temperature	0.08	(-0.02, 0.17)	0.13	(0.04, 0.22)
Year of the survey				
2	-0.98	(-1.33, -0.67)	-0.13	(-1.32, 0.8)
3	-0.74	(-2.51, 1.01)	-0.02	(-1.81, 1.48)
Variances				
Spatial $(\sigma_{\phi}^{(D)^2})$	0.94	(0.57, 1.56)	0.87	(0.52, 1.46)
Temporal $(\sigma_{\varepsilon}^{(D)^2})$	0.82	(0.46, 1.45)	0.88	(0.53, 1.52)
Nugget $(\sigma_{e}^{(D)^2})$	1.02	(0.75, 1.30)	0.87	(0.61, 1.19)
Range (in km)	38.8	(22.4, 51.0)	38.8	(22.4, 51.0)
Dispersion parameter (r)	0.98	(0.74, 1.17)	0.59	(0.51, 0.70)

For *An. gambiae*, distance to water bodies, NDVI, day temperature and night temperature were associated with mosquito density. Higher day temperatures and longer distances from breeding sources were associated with decline in mosquito density. An increase in vegetation led to an increase in mosquito abundance. Spatial, temporal and non-spatial variances were almost equal.

Over-dispersion was present only for *An. gambiae* (r = 0.6, (95% CI: 0.5, 0.7)). The minimum distance at which the spatial correlation was below 5% was 39 km (95% CI: 22.4 km, 51 km) for both species.

3.3.5 Model-based results: Sporozoite rate data

Lag analysis shows that *An. funestus* SR was related to total rainfall of the survey month, average NDVI, average night temperature for the two months preceding the survey, and average day temperature of current and previous month. Similarly, *An. gambiae* SR was driven by the average NDVI of the survey month; total rainfall, and average (of day and night) LST of the current and previous month. Results of SR rate models with spatial and temporal random effects were presented because they provided the best performance with a predictive ability of 40% of the test locations within a 95% Bayesian credible interval. Parameter estimates of the geostatistical logistic regression models are shown in Table 3.3.

Table 3. 3: Multivariate spatio-temporal analysis for sporozoite rate

Parameters	*An. funestus*		*An. gambiae*	
	Co-efficients		Co-efficients	
	Median	95% CI	Median	95% CI
Intercept	-0.83	(-2.74, 0.65)	-1.75	(-4.73, 0.04)
Altitude	0.01	(0.002, 0.02)	0.01	(0.00, 0.02)
Distance to water bodies	-0.22	(-0.33, -0.11)	-0.06	(-0.15, 0.06)
NDVI	-0.6	(-1.30, 0.03)	-0.89	(-1.97, 0.14)
Rainfall	-0.001	(-0.01, 0.004)	-0.01	(-0.01, 0.00)
Season(Wet)	-0.1	(-0.25, 0.06)	0.36	(-0.18, 1.06)
Day temperature	-0.01	(-0.05, 0.04)	-	-
Night temperature	-0.12	(-0.21, -0.03)	-	-
Average temperature	-	-	-0.07	(-0.14, 0.04)
Year of the survey				
2	-0.78	(-1.73, 0.21)	-0.97	(-2.26, 0.17)
3	-0.62	(-1.8, 0.44)	-0.48	(-1.49, 0.28)
Variances				
Spatial ($\sigma_{\phi}^{(S)^2}$)	0.6	(0.40, 0.96)	0.77	(0.56, 1.09)
Temporal ($\sigma_{\varepsilon}^{(S)^2}$)	0.3	(0.15, 0.63)	0.38	(0.18, 0.88)
Range (in km)	4.1	(2.0, 9.2)	2.0	(1.0, 4.1)
Mixing proportion (π)	0.54	(0.53, 0.56)	0.54	(0.53, 0.55)

Altitude, distance to the nearest water bodies and night temperature were associated with *An. funestus* SR. Higher night temperatures were associated with low SR in that area. Similarly, places closer to water bodies observed a higher proportion of infectious mosquitoes than others. A positive association between *An. funestus* SR and altitude was estimated. Spatial variability ($\sigma_{\phi}^{(S)^2}$ =0.6, (95% CI: 0.4, 1.0)) was higher than temporal one ($\sigma_{\varepsilon}^{(S)2}$ =0.3, (95% CI: 0.2, 0.6)).

The minimum distance at which the spatial correlation is below 5% was 4.1 km (95% CI: 2.0 km, 9.2 km). On the other hand, altitude was the only factor associated with SR for *An. gambiae*. Spatial variation from *An. gambiae* sporozoite model ($\sigma_{\phi}^{(S)2}$ =0.8, (95% CI: 0.6, 1.1)) was twice

as high as the temporal one ($\sigma_{\varepsilon}^{(S)2}$ =0.4, (95% CI: 0.2, 0.9)). The minimum distance at which the spatial correlation is below 5% was 2.0 km (95% CI: 1 km, 4 km). This shows a slower decay of the correlation with distance for the *An. funestus* SR compared with *An. gambiae.*

3.3.6 Model-based results: EIR estimates

Figure 3 shows the temporal patterns in the EIR values that were captured by the spatio-temporal models. Smooth monthly EIR maps (Figure 4) clearly show a seasonal pattern, ranging from almost no infective bites in the dry season to the highest number of infective bites toward the end of wet season. It is evident from the maps that areas close to water bodies experienced high EIR.

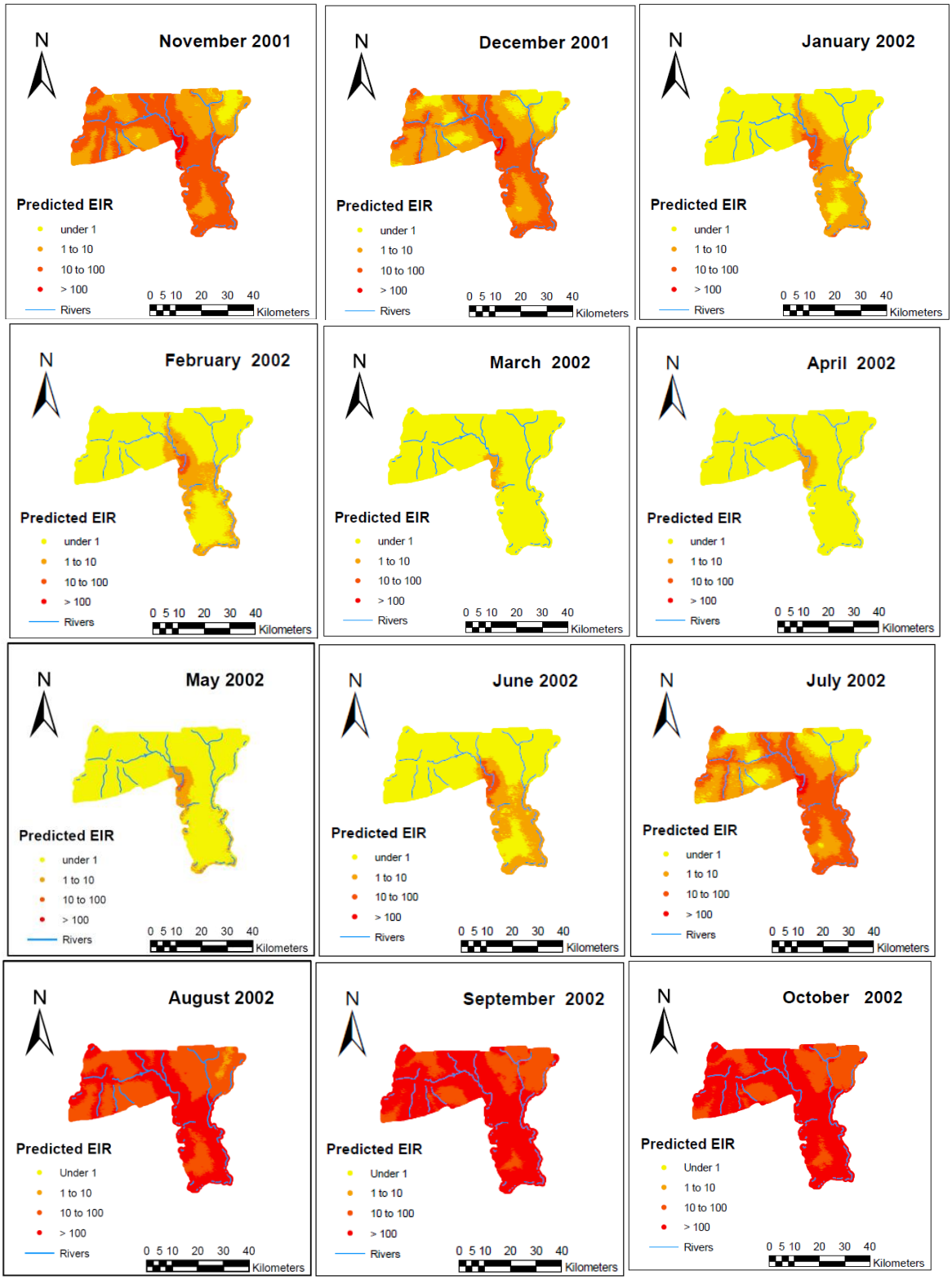


Figure 3. 4: Predicted EIR by month for the first year

3.4 Discussion

This is the first study assessing malaria transmission heterogeneity in the Navrongo HDSS using a comprehensive entomological dataset and rigorous geostatistical and temporal models, which take into account data characteristics. These data indicate the presence of seasonal, spatial and year-to-year variation within a small geographical area (1,675 km^2) in northern Ghana. The findings confirm previous studies reporting heterogeneity in malaria transmission in small areas. In particular, spatio-temporal variation has been observed in coastal Kenya(Mbogo et al., 2003), in Kilombero valley in Tanzania (Drakeley et al., 2003), in some selected Ugandan villages(Okello et al., 2006) and in a low transmission zone in Sudan(Hamad et al., 2002).

Transmission in the Kassena-Nankana district is high (EIR > 100 ib/p/y) especially during the wet season. An entomological survey conducted in the same district between June 2001 and May 2002 recorded EIR of 630 ib/p/y in the irrigated zone within the southern part of the district (Appawu et al., 2004) which is lower than the one observed in the first year of the MTIMBA project. The drop of EIR after the first year may be explained by variations in laboratory testing. The ELISA tests for the first year were carried out in a different laboratory from those in the remaining two years, making it possible that inter-laboratory differences contribute to inter-annual variation. The year effect included in the model is, therefore, aliased with any laboratory differences. Consequently, there will be more confidence in EIR comparisons between locations than those that depend on inter-annual differences.

This study confirmed the presence of *An. funestus* and *An. gambiae* malaria vector species in the region(Abonuusum et al., 2010; Appawu et al., 2004; Dery et al., 2010), with both acting as major vectors. NDVI, distance to water bodies and temperature were associated with mosquito

density for both species. Compounds located close to water bodies were more likely to have high mosquito densities. The Kassena-Nankana district in northern Ghana has many irrigation dams that were constructed to increase food production in the area. There are also many small dugout reservoirs in the area which supply water to various communities especially in the dry season (Binka et al., 1998; Owusu-Agyei et al., 2002). These water bodies can be favourable breeding grounds and responsible for mosquito abundance in neighbouring compounds. The data showed that a reduction in day temperature favoured higher number of mosquito bites in the area. The NHDSS where data were collected experiences high temperatures in some months (18 °C to 45 °C). Temperatures close to 40 °C reduce mosquito survival, hence their density(Craig et al., 1999) . Although rainfall had a positive relationship with mosquito density, the association was not statistically important. However, rainfall is known to have a direct relationship with other factors, such as vegetation, that were found to positively influence mosquito abundance. A positive correlation between precipitation and mosquito density for both *An. funestus* and *An. gambiae* has already been observed in other places.

A seasonal pattern in mosquito density was observed for both species. High mosquito densities were observed in the rainy season for all the three years and low densities during the dry season. However, SR was higher in the dry than the rainy season during the second and third year. In addition, *An. gambiae* SR in the dry period were higher than that of *An. funestus* for the entire survey period. There was no evidence of variations in SR between species in the rainy season. More infected mosquitoes during dry seasons have already been observed in other areas (Charlwood et al., 1995). This implies that most surviving adult mosquitoes in dry seasons are likely to be infectious.

The shortest distance at which the spatial correlation was below 5% was lower for SR than mosquito densities, suggesting that SR depends largely on local conditions rather than environmental factors. On the other hand, mosquito densities had strong spatial correlation and therefore they are more likely to be driven by environmental factors, especially vegetation which was the major predictor in the Navrongo area. Climate and environmental factors influence malaria transmission and its effects. In this district, malaria illnesses and mortality are observed thought the year with peaks in the wet season (Binka et al., 1994; Koram et al., 2000). Blood transfusion, especially in young children, due to anaemia is more common in the rainy season (Owusu-Agyei et al., 2002) .

The EIR maps clearly depict spatial heterogeneity despite the relative small size of the HDSS. The high EIR estimate in the southern part, which is mainly covered by irrigation dams, has been reported previously (Appawu et al., 2004). Even during the dry season, transmission in the area remained high. In addition, the geographical pattern of EIR was similar across the three years of the project. The spatial and temporal variances of the mosquito density data accounted for about 33% each out of the total variation. However, SR data explained 67% and 33% of the total variation, suggesting that spatial heterogeneity was twice as high as the temporal one. Although space-time heterogeneity could explain total variation of the SR data, there was a remaining 34% unexplained variation for the densities. In principle, focussed malaria control conducted in the knowledge of these patterns of variation might be more effective than generalized intervention programmes, but no intervention programme is likely to be able to adapt to variations on this scale.

Bayesian geostatistical models are the state-of-art methodology to analyse space and time heterogeneity in malaria transmission and have been used to assess malaria risk using prevalence data(Ashton et al., 2011; Gemperli et al., 2006; Giardina et al., 2012; Hay et al., 2009; Riedel et al., 2010). However, entomological data have large number of zeros, which cannot be estimated by standard geostatistical models. In particular, the Navrongo data had 44% and 72% of locations with zeros for density and SR, respectively. Entomological data were sparse in the other two MTIMBA sites (i e, Rufiji and Kisumu). This problem was addressed by developing geostatistical zero-inflated formulations of binomial models (GZIB) for analysing SR(Amek et al., 2011). Zero-inflated analogues of negative binomial models (Amek et al., 2012; Rumisha, 2013) were also applied to take into account excess zeros in the density data. These models were able to improve EIR predictions obtained from standard geostatistical analogues.

The EIR estimates of this study will be used further to analyse the relationship between malaria transmission intensity and mortality as part of the ongoing work for the MTIMBA project.

3.5 Appendix 2

Spatio-temporal modelling of sporozoite rate

Sporozoite rate are binomial data modelled via logistic regression. We assumed that for a specific species, the number of positive mosquitoes $Y_{it}^{(S)}$ out of all tested (N_{it}) follows a binomial distribution; that is $Y_{it}^{(S)} \sim Bin\left(N_{it,} p_{it}\right)$ with parameter p_{it}, the sporozoite rate at location i and time t. To account for spatial and temporal variation in the data, we introduced compound and monthly random effects $\phi_i^{(S)}$ and $\varepsilon_t^{(S)}$ respectively. These were modelled together with covariates X_{it} on a logit scale as $\log it(p_{it}) = X_{it}{}^T \beta^{(S)} + \phi_i^{(S)} + \varepsilon_t^{(S)} + e_i^S$, where $\beta^{(S)} = \left(\beta_1, \beta_2, ..., \beta_k\right)^T$ is a vector of regression coefficients. The spatial random effects $(\phi_i^{(1)})$ are assumed to be latent observations from a Gaussian spatial process with zero mean and variance-covariance matrix $\Sigma^{(S)}$, where $\phi_i^{(S)} = \left(\phi_1, \phi_2, ..., \phi_n\right)^T \sim N(0, \Sigma^{(S)})$. The covariance between any particular pair of locations was considered to be a function of distance between the locations, that is $\Sigma_{ij}^{(S)} = \sigma_\phi^{(S)^2} \exp\left(-d_{ij}^{(S)} \rho^{(S)}\right)$ where $d_{ij}^{(S)}$ is the Euclidean distance between locations i and j, $\sigma_\phi^{(S)^2}$ is the spatial variance and $\rho^{(S)}$ is the correlation decay parameter with a range defined as $3/\rho^{(S)}$. Temporal random effects were modelled by a first order autoregressive process [AR (1)] with variance $\sigma_\varepsilon^{(S)^2}$ which allows correlation between consecutive time periods (J. L. Hay and Pettitt, 2001a). Non spatial random effects $e_i^{(S)}$ with zero mean and variance $\sigma_e^{(S)^2}$ were added to the model in order to account for unexplained variation in the data. Seasonality variable dry/wet was included in the model also as a covariate.

Spatio-temporal modelling of mosquito density

Let the number of mosquitoes $Y_{it}^{(D)}$ caught at location i and time t follows a negative binomial distribution, $Y_{it}^{(D)} \sim NB\left(\mu_{it}, r\right)$, with μ_{it} the mean and r, the over-dispersion

parameter(Venables and Ripley, 2003). To account for spatial and temporal variation in the data, we introduced location and monthly random effects $\phi_i^{(D)}$ and $\varepsilon_t^{(D)}$ respectively. Non spatial random effects $e_i^{(D)}$ with zero mean and variance $\sigma_e^{(D)^2}$ were added to the model in order to account for unexplained variation in the data. The relationship between mean density of each species (μ_{it}), the covariates X_{it} and the random effects is modelled as;

$\log(\mu_{it}) = X_{it}{}^T \beta^{(D)} + \phi_i^{(D)} + \varepsilon_t^{(D)} + e_i^{(D)}$ where, X_{it} is the vector of covariates at location i for time t and $\beta^{(D)} = (\beta_1, \beta_2, ..., \beta_k)^T$, the vector of regression coefficients. Like seasonality, spatial and temporal random effects are defined and modelled in similar way as described in the spatio-temporal model of sporozoite rate above.

Zero inflated models

Our entomological data had many locations with either no mosquitoes (56%) or uninfected mosquitoes (72%). This calls for zero inflated models that add extra weight to the probability of observing zero (Lambert, 1992; Vounatsou et al., 2009). Such models have two components; one arising from either binomial or negative binomial distribution and another for excess zero that cannot be estimated by the model. In such models, π_{it} is the mixing proportion and the corresponding $(1-\pi_{it})$ is the probability of observing an outcome arising from either binomial or negative binomial distribution. The model therefore is written as follows;

$$P(Y_{it} = y_{it}) = \begin{cases} \pi_{it} + (1-\pi_{it}) f(y_{it}), & y_{it} = 0 \\ (1-\pi_{it}) f(y_{it}), & y_{it} > 0 \end{cases}$$

where $f(y_{it})$ is the binomial or negative binomial density function depending on the outcome data.

Models were fitted assuming that either the mixing proportion is constant throughout space and time i.e $\log it(\pi_{it}) = \beta_0^{(m)}$ or that it is a function of environmental covariates (NDVI/EVI, LST, and rainfall); i.e $\log it(\pi_{it}) = X_{it}^{T}\beta^{(m)}$, where X_{it}^{T} is the set of covariates and $\beta^{(m)} = (\beta_0, \beta_1, \beta_2, \beta_3)^T$ is the vector of regression coefficients for each mosquito species. Based on model validation, SR data were fitted with constant value, while density with a mixing proportion derived from a function of covariates.

Model fit and implementation

We applied Bayesian inference by combining likelihood function and prior distributions to form the posterior distribution that was used in estimating model parameters. Prior distributions specified for all model parameters were; normal with zero mean and large variance for regression coefficients $[(\beta^{(1)}, \beta^{(D)}, \beta^{(S)});\ \beta \sim N(0, 10^2)]$, inverse gamma for the variances $[(\sigma_\phi^2, \sigma_\varepsilon^2, \sigma_e^2);\ \sigma^2 \sim IG(2.01, 1.01)]$, gamma for the dispersion parameter $[r \sim G(0.01, 0.01)]$, uniform for decay parameter $[\rho \sim U(a\ , b\ )]$ and beta distribution for the mixing proportion $[\pi \sim Be(1,1)]$.

Markov Chain Monte Carol (MCMC) simulation algorithm was used in estimating the model parameters. We used a single chain sampler of 250000 iterations with an initial burn-in of 10000 OpenBUGS version 3.1.1 software was used for parameter estimation. FORTRAN program written by the authors was used for Bayesian Kriging in order to predict both SR and density at locations where data were not collected (Diggle et al., 1998).

Space time prediction of EIR

Model-based products for SR and density at high resolution were combined to generate EIR. These models were also used to predict EIR at unobserved locations over a grid of 31308 pixels at a $250m^2$ spatial resolution.

Chapter 4: Relationship between all-cause mortality and entomological inoculation rate in Navrongo Demographic surveillance site, Ghana

Authors: Simon Kasasa[1,2,3], Thomas Smith[1,2], Victor Asoala[4], Laura Gosoniu[1,2] , Francis Anto[5], Martin Adjuik[6], Cletus Tindana[4], Seth Owusu-Agyei[7,] Penelope Vounatsou[1,2 *]

[1]Swiss Tropical and Public Health Institute, Socinstrasse 57, P.O. Box 4002 Basel, Switzerland, Tel: +41612848109; Fax: +41612848105
[2]University of Basel, Switzerland
[3]School of Public Health, Makerere University college of Health Sciences, Uganda
[4]Navrongo Health Research Centre, Ghana
[5]School of Public Health, University of Ghana, Legon, Ghana
[6]INDEPTH Network Secretariat, Accra, Ghana
[7]Kintampo Health Research centre, Ghana Health Services, Ministry of Health

*Corresponding Author

This manuscript is prepared for submission to Trends in Parasitology Journal

Abstract

Health and Demographic Surveillance Systems (HDSS) provide reliable mortality data for measuring progress towards Millennium Development Goals targeting malaria. The disease has overburdened many Africans making its control one of the highest priorities. Understanding this relationship will guide malaria control programmes about required efforts to reduce transmission to levels where the disease will no longer be of major public health importance. The INDEPTH network integrated the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project into routine activities of selected HDSS in Africa. Entomological data were collected bi-weekly and mortality data were extracted from HDSS databases. In this work, we assessed the mortality-malaria transmission relationship using data collected from Navrongo HDSS between 2001 and 2004. Spatio-temporal logistic regression models were fitted to obtain age-specific mortality risk estimates. The model considered 3 covariates; entomological inoculation rates (EIR) estimates with their measurement errors, age and insect treated nets (ITN) ownership (for children <5 years). Model parameters were used to estimate excess mortality at different EIR levels. A total of 5412 deaths were registered with an overall mortality rate (MR) of 14 per 1000 person-years of observation. The infants (0-11 months) experienced the largest risk of dying followed by elderly people. The annual mortality rates declined consistently in all age groups. The overall mortality for male (MR=16) was higher than that of female (MR=13). The increase in natural logarithm transformed EIR were positively associated with all-cause mortality at all age categories namely; neonates (2%), post-neonates (12%), children (13%), school age children (1%), adults of 15 to 29 years (2%), adults aged 30 and 50 years (1%), and elderly (8%). Despite the largest hazard ratio among individuals between 1 and 4 years, the excess mortality in this age group was lower than that of infant, a situation similar to what was observed in Kisumu and Manhica HDSS. There was a positive trend in the magnitude of hazard ratios with age among the under-fives and a decline between 15 and 59 years. Possession of nets offered a protective effect among all children <5 years.

4.1 Introduction

More than half a million people die each year in Sub-Saharan Africa (SSA) due to malaria exposure (WHO Report, 2011). The disease is clustered among the under-five and pregnant women (Lusingu et al., 2004; Rowe et al., 2006; Steketee et al., 2001). However, the relationship between mortality and transmission intensity is not clear. It has been assumed that malaria control programs aimed at reducing transmission intensity in endemic areas might delay the acquisition of immunity and hence shift the disease burden to an older population (Snow and Marsh, 1995).

The entomological inoculation rate (EIR), defined as the number of infective mosquitoes bites per person per unit time interval is a direct measurement of human exposure to malaria infection in endemic areas (Beier et al., 1999). Snow et al., (Snow and Marsh, 1995), conducted a meta-analysis and observed no clear relationship between all-cause mortality and malaria transmission intensity. However, previous studies had observed reductions in severe morbidity and mortality after interrupting malaria transmission using insect-treated bed nets in different countries (Alonso et al., 1991; Binka et al., 1996; D'Alessandro et al., 1995; Nevill et al., 1996). Further analyses using updated data showed a positive linear association especially among the infants (Ross and Smith, 2006; Smith et al., 2001) but no clear trend was observed among children (1-4 years) (Smith et al., 2001). However most of findings were based on historical data collected at different time points and the focus was mainly on children under five years. This is partly due to lack of reliable mortality data in malaria endemic countries caused by their weak health systems and poor civic registrations processes (Mathers et al., 2005; Ruzicka and Lopez, 1990).

The Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project was implemented by the INDEPTH network in selected Health and Demographic Surveillance

Systems (HDSS) in SSA. The aim of the MTIMBA project is to assess the mortality-malaria transmission relationship.

Data from the three MTIMBA sites of Rufiji in Tanzania (Rumisha, 2013), Kisumu in Kenya (Amek, 2013) and Manhica in Mozambique (Kasasa, in preparation) reported a positive log-linear relationship between all-cause mortality and transmission intensity among the under-fives. However there were differences in the magnitude and direction of mortality risk in other age groups. It is imperative that more MTIMBA data site-specific analyses are carried out before drawing conclusions for malaria control programs.

We extended the MTIMBA work by analyzing the space-time relationship between mortality and malaria transmission intensity in different age groups using data collected from the Navrongo HDSS between November 2001 and October 2004. Mortality data were linked to EIR estimates generated from our previous work for the same area (Kasasa et al.,2013). Mortality risks were compared by gender, social economic status and ITN ownership.

4.2 Methods and materials

4.2.1 Study site

The Navrongo health and demographic surveillance system (NHDSS) evolved from the Ghana vitamin-A supplementation trial of Kassena-Nankana that started in 1989 (Ghana VAST Study Team, 1993). The Navrongo Health Research Centre (NHRC) established the HDSS in 1993 to continue monitoring health and demographic outcomes in Ghana's rural Savahan zone. The NHDSS is located in the administrative district of Kassena-Nankana, in northern Ghana bordering Burkina Faso. The district covers a land surface area of 1,675 square kilometres. The 1999 population was close to 140,000 people and the majority is for subsistence farmers. The population comprises of two distinct ethnolinguistic groups; the Kassena and the Nankani. There

are two distinct seasons; the wet, between April and October and a dry period that covers remaining months of the year. Malaria is endemic and *Plasmodium falciparum* is transmitted by both *An. gambiae* and *An. funestus*. Kassena-Nankana district has one hospital and four health centres serving the entire area. Additional characteristics of the district and the HDSS have been documented elsewhere (Appawu et al., 2004; Binka et al., 1999; Owusu-Agyei et al., 2007).

4.2.2 Malaria transmission and mortality data

Mosquitoes were collected from 2803 randomly selected geo-referenced compounds in the NHDSS over a period of three years. . Entomological data were collected every fortnight, while mortality data were extracted from HDSS databases for the three years MTIMBA period. Independent Bayesian geostatstical zero-inflated models for density (negative binomial) and sporozoite rate (logistic regression) were fitted separately taking into account climate and environmental factors. Bayesian Kriging was used to predict sporozoite rate and mosquito density at unsampled locations. EIR estimates were generated at monthly interval by multiplying sporozoite rates and man biting rates. The latter were derived from density data after adjusting for a collection bias between light traps and human landing catches methods (Lines et al., 1991). Monthly malaria exposure surfaces at high spatial resolution were estimated for the entire HDSS as described in our previous work (Kasasa et al., 2013).

Population demographic characteristics were extracted from routinely collected data in the HDSS for the entire MTIMBA period. Deaths, births and migrations are some of the key outcomes that are routinely monitored. Personal information such as gender and age were also extracted. All the 13266 geo-referenced compound where mortality was monitored between November 2001 and October 2004 were linked to the nearest predicted EIR.

4.2.3 Socio-economic and intervention data

The NHDSS collected household assets data once during the MTIMBA project period. We used principal components analysis (PCA) to construct a socio-economic status (SES) index for each household (Vyas and Kumaranayake, 2006). The household assets used were motorbikes, bicycles, radio, sewing machines, and livestock (cattle, sheep, goat and guinea fowl). The SES index was categorized into quintiles to rank households from poorest (first quintile) to least poor (fifth quintile). The number of nets in every compound was also collected together with other assets. In this analysis, presence of atleast a net in a compound was used as a proxy measure of ITN use by all members and this represented the malaria intervention data from the HDSS.

4.2.4 Statistical analysis

The NHDSS population was stratified into different age groups namely; neonates (0-28 days), post-neonates (1-11 months), children (1-4 years), school-age (5-14 years), adults (15-59 years) and old age (60 years and above). Separate analyses were carried out for each age group. Crude mortality rates were expressed per 1000 person-years at risk (person-years).

Survival models were approximated via logistic regression, treating time to death as discrete (Allison, 1982; Singer and Willet, 1993) with dependence on time-dependant covariates. Malaria transmission intensity varies with time partly due to changes in climate and environmental factors. Time to death data were therefore split at monthly intervals in order to generate Bernoulli and binomial data that were modelled via logistic regression formulations. Bayesian geostatistical spatio-temporal models were fitted to allow for spatial and temporal correlations in the data. Spatial correlation was modelled via village-specific random effects which were assumed to be latent observations from a multivariate Gaussian spatial process with a zero mean. The covariance of the process assumed an exponential correlation function of

distance between any pair of villages. Temporal random effects were modelled by the first order autoregressive processes (Hay and Pettitt, 2001). Any remaining non-spatial variation (nugget parameter) was considered by additional village random effects which are independent and normally distributed with zero mean. Mortality was related to natural logarithm transformed EIR of the previous month. Predicted EIR was incorporated in the model as a covariate with measurement error to account for the prediction uncertainty (Gemperli, 2003).

All the analyses were implemented in OpenBUGS statistical software (Lunn et al., 2009). Parameters were estimated using Markov Chain Monte Carol (MCMC) simulation algorithm (Gelfand and Smith, 1990). Details of mathematical description for all models used are given in appendix 3

4.2.5 Excess Mortality due to malaria exposure

The malaria attributed mortality was computed by considering the difference between mortality rate when transmission intensity is more than zero and at zero EIR and was expressed per 1000 person-years. The probabilities of dying were transformed into rates and the Taylor series approximation was used to generate mortality at zero EIR. Details of the models used in computing excess mortality have been described and documented in the analyses of Rufiji and Kisumu HDSS data (Amek, 2013; Rumisha, 2013). Final survival model coefficients for each age group were used to compute the probability of death via a logistic regression model over a range of EIR between 0.1 and 300 infectious bites per person per month.

4.3 Results

4.3.1 Description of mortality data

A total of 163128 individuals from 33 villages were included in the analysis. There were more females (53%) in the NHDSS. Majority of the people (53%) were between 15 and 60 years, followed by children under 15 years (38%). During the 3 years follow-up period, 5412 deaths occurred with 384748.4 person-years at risk were registered leading to an overall mortality rate of 14.1 (95% CI: 13.6, 14.4) per 1000 person-years. The highest number of deaths occurred among the elderly (36%), followed by those between 15 and 60 years (33%) and the least number was observed among the school going children (6%). More than half (52%) of the compounds had atleast a mosquito net.

Figure 4.1 clearly shows crude mortality rates by age groups over the 3 year period. . The risk of dying was higher among the infants and declined from 87 to 72 per 1000 person-years between the first and third year of life. In this population, crude mortality estimates among the elderly (at least 60 years) are comparable with those of children who had not attained their first birthdays.

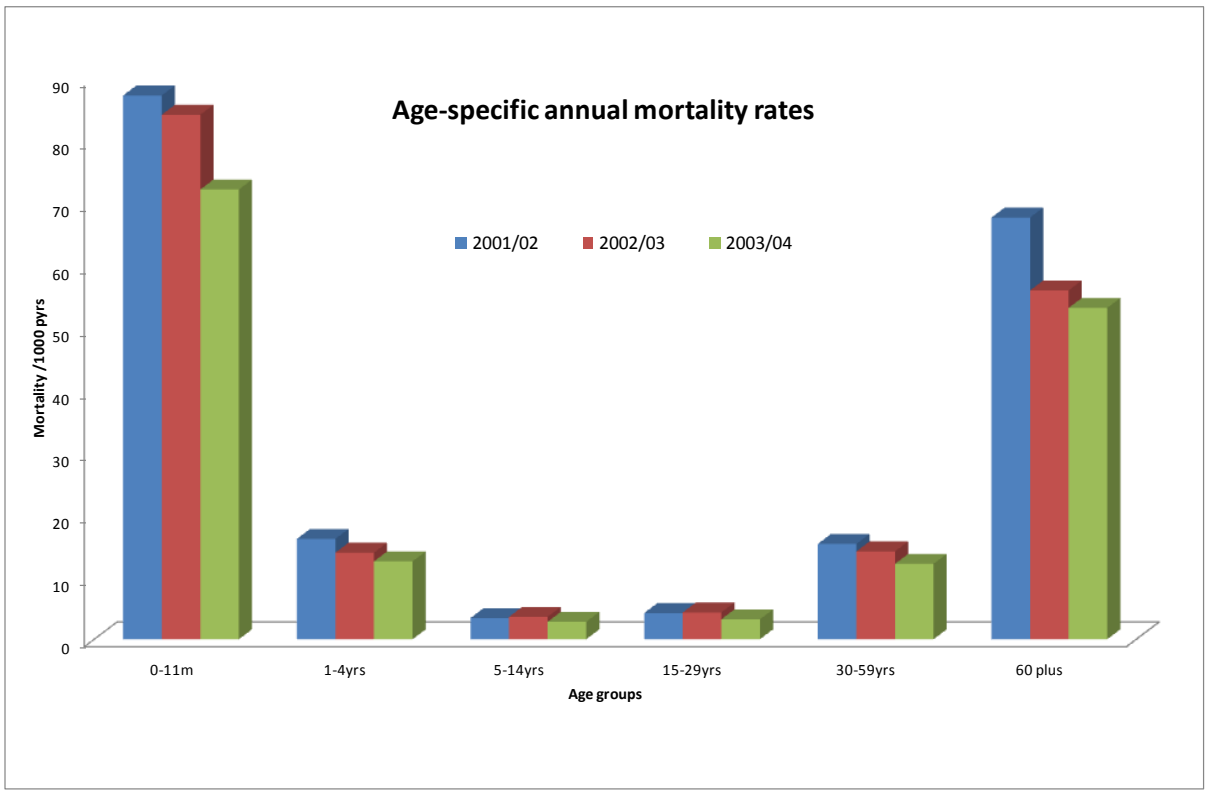


Figure 4. 1: Age specific annual all-cause mortality rates

Mortality rates for the old population were 68, 56 and 53 per 1000 person-years for the first to third year respectively. Mortality rates declined consistently in all the age groups by 17% in infants, 22% (children 1-4 years), 18% (school age), 24% (adults 15-29 years), 21% (adults 30-59 years) and 21% in old people. As indicated in figure 4.1, the risk of dying was least among people aged between 5 and 30 years. Over the three years period, the highest risk of death was observed among children below one year (80.6/1000 person-years) and lowest from those who had reached their fourth birthday (5.1/1000 person-years). The overall under five mortality rates of 31, 28 and 25 per 1000 person-years were observed in the HDSS from years 1 to 3 respectively. Mortality rates among the under fives declined with increase in age.

Table 4.1 shows mortality patterns across age groups in relation to SES, net ownership and gender. The crude mortality rates show that male (MR=15.6;[95 CI; 15.0, 16.2]) had a higher

risk of dying than female (MR=12.7;[95 CI; 12.2, 13.2]). Apart from the elderly (60 years plus) population, the risk of death was lower for individuals who resided in households that owned at least one net. Analyses by SES index indicated that mortality rates for the infants increased from the first to the second quintile followed by a consistent decline up to the fifth quintile (least poor). There was a decline in overall crude mortality rates from poor to least poor families of 14.5, 13.7 and 13.2 per 1000 person-years respectively. The poorest to the least poor ratio of 1.23 indicates some variations in mortality risk among the different families. The ratios (poorest/least poor) among different age groups ranged between 0.65 and 1.31 as indicated in Table 4.1.

Table 4. 1: All-cause mortality rates per 1000 person-years

Variable	0-11months	1-4 yrs	(5 - 14 yrs)	(15-59 yrs)	60yrs plus	All
SES	MR (95% CI)	MR (95% CI)	MR (95% CI)	MR (95% CI)	MR (95% CI)	MR (95% CI)
Poorest	90.0 (71.8, 113)	14.2 (10.5, 18.7)	2.2 (1.4, 3.4)	11.4 (10.0, 13.0)	62.8 (55.1, 71.3)	16.3 (15.1, 17.7)
Poorer	93.3 (78.2, 111.3)	13.7 (10.8, 17.3)	3.3 (2.4, 4.4)	9.0 (8.0, 10.2)	62.2 (55.1, 70.1)	14.3 (13.3, 15.3)
Poor	89.0 (76.3, 103.2)	10.4 (11.5, 16.9)	3.1 (2.3, 4.0)	8.7 (7.8, 9.6)	63.6 (57.6, 70.1)	14.5 (13.7, 15.4)
Less poor	78.7 (68.8, 89.7)	15.3 (13.1, 17.8)	3.3 (2.6, 4.0)	8.7 (7.9, 9.5)	54.5 (49.7, 60.0)	13.7 (13.0, 14.4)
Least poor	68.5 (59.5, 78.5)	12.5 (10.6, 14.7)	3.4 (2.8, 4.1)	8.7 (7.9, 9.4)	55.8 (51.0, 60.8)	13.2 (12.5, 13.9)
Ratio (Q1/Q5)	1.31	1.14	0.65	1.31	1.13	1.23
ITN						
No	92.7 (84.0, 102.1)	14.9 (13.2, 16.8)	3.2 (2.7, 3.8)	9.7 (9.1, 10.4)	56.8 (53.0, 60.8)	14.8 (14.2, 15.4)
Yes	71.2 (64.4, 78.5)	13.1 (11.6, 14.7)	3.2 (2.7, 3.7)	8.5 (8.0, 9.1)	60.3 (56.7, 64.0)	13.5 (13.0, 14.0)
Sex						
Female	80.3 (72.6, 88.6)	13.5 (11.9, 15.3)	2.4 (2.0, 2.9)	7.6 (7.0, 8.0)	52.3 (49.1, 55.8)	12.7 (12.2, 13.2)
Male	81.1 (73.5, 89.2)	14.3 (12.7, 16.1)	3.9 (3.4, 4.5)	10.8 (10.1, 11.5)	66.8 (62.7, 71.1)	15.6 (15.0, 16.2)

MR: mortality rate, Q1- poorest; Q5-least poor

4.3.2 Model-based results

The parameter estimates of the spatio-temporal model in Table 4.2 indicate a positive natural logarithmic relationship between malaria exposure and all-cause mortality across all age groups, namely the neonates (hazard ratio; HR 1.02;[95% CI: 0.09, 1.15]), post-neonates (HR 1.12;[95% CI: 0.98, 1.26]), children (HR 1.13;[95% CI: 1.00, 1.27]), school age children (HR 1.01;[95% CI: 0.94, 1.07]), adults of 15 to 29 years (HR 1.02;[95% CI: 0.97, 1.06]), adults aged 30 and 50 years (HR 1.01;[95% CI: 0.98, 1.05]), and elderly (HR 1.08;[95% CI: 1.02, 1.06]). The hazard ratios refer to the effect of an e-fold change in EIR. For instance, an e-fold increase in EIR was associated with increased mortality risk of 2%, 12% and 13% among neonates, post-neonates and children respectively.

Table 4. 2: Spatio-temporal multivariate posterior estimates for all-cause mortality

Parameters	0-28days	1-11months	1-4 yrs	(5 - 14 yrs)	(15-29 yrs)	(30-59 yrs)	60yrs plus
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Constant	0.13 (0.09, 0.22)	0.01 (0.003, 0.01)	0.004 (0.002, 0.01)	-	-	-	-
Log EIR *	1.02 (0.90, 1.15)	1.12 (0.98, 1.26)	1.13 (1.00, 1.27)	1.01 (0.94,1.07)	1.02 (0.97,1.06)	1.01 (0.98, 1.05)	1.08 (1.02,1.16)
Age **	0.85 (0.84, 0.86)	0.91 (0.89, 0.94)	0.50 (0.45, 0.54)	0.52 (0.41,0.64)	1.22 (0.98,1.51)	1.69 (1.52,1.88)	1.06 (1.05,1.06)
Variances							
Spatial (σ_{ϕ}^2)	0.20 (0.10, 0.44)	0.19 (0.09, 0.41)	0.16 (0.09, 0.34)	0.18 (0.09,0.36)	0.17 (0.09,0.34)	0.13 (0.07,0.27)	0.13 (0.07, 0.28)
Temporal (σ_{ε}^2)	0.14 (0.08, 0.26)	0.14 (0.08, 0.25)	0.16 (0.09, 0.29)	0.17 (0.09,0.31)	0.15 (0.08,0.27)	0.10 (0.06, 0.17)	0.09 (0.06, 0.16)
Nugget (σ_{e}^2)	0.17 (0.09, 0.35)	0.15 (0.09, 0.30)	0.14 (0.08, 0.27)	0.17 (0.09,0.32)	0.16 (0.09,0.33)	0.12 (0.07, 0.22)	0.12 (0.07, 0.22)
Range (in km)	5.77 (0.78, 45,85)	9.33 (0.80, 47.1)	5.54 (0.79, 46.62)	1.50 (0.84, 21.3)	1.73 (0.84, 26.37)	2.48 (0.84, 49.05)	3.32 (0.85, 50.58)

yrs -years ; * Natural logarithim of EIR; ** Units of age were based on model category (days, months and yrs)

The relationship between mortality and malaria transmission intensity was statistically important among children (1-4 years) and old age people (60 years and above). Individuals in the former age group experienced the highest risk of dying 13% that was associated with increase in

transmission intensity. Age was negatively associated with mortality for all the age categories below 15 years.

The under five sub-groups' risk of dying due to malaria exposure increased with increase in age namely; 2% (neonates), 12% (post-neonates) and 13% (children). The minimum distance at which the spatial correlation is less than 5% ranged between 2km and 9km across all age categories suggesting a geographical dependency in mortality exposures. Spatial and temporal variations were statistically related to mortality in all age groups.

Table 4.3 shows parameter estimates from Bayesian geostatistical model where ITN was included as a covariate. Only results for the under fives are shown. The analysis shows a positive log-linear relationship between mortality and malaria transmission intensity among the three age groups namely; neonates (HR 1.02;[95% CI: 0.90, 1.15]), post-neonates (HR 1.13;[95% CI: 1.00, 1.27]), and children between 1 and 4 years (HR 1.13;[95% CI: 0.99, 1.27]). The hazard ratios associated with EIR for post-neonates and children (1-4 years) were equal (13%). Age was negatively associated with a reduction in mortality.

Table 4. 3: Posterior estimates for all-cause under five mortality with ITN

Parameters	*Neonates (0-28 day)* HR (95% CI)	*Post-neonatal (1-11 months)* HR (95% CI)	Child (1-4 years) HR (95% CI)
Constant	0.14 (0.09, 0.24)	0.006 (0.004,0.010)	0.004 (0.003, 0.30)
EIR (Log e scale)	1.02 (0.90, 1.15)	1.13 (1.00, 1.27)	1.13 (0.99, 1.27)
ITN	0.85 (0.67, 1.09)	0.79 (0.66, 0.95)	0.50 (0.45, 0.54)
Age	0.85 (0.84, 0.86)	0.91 (0.89, 0.93)	0.92 (0.76, 1.10)
Variances			
Spatial (σ_{ϕ}^2)	0.20 (0.10, 0.43)	0.18 (0.09, 0.39)	0.16 (0.09, 0.33)
Temporal (σ_{ε}^2)	0.14 (0.08,0.26)	0.14 (0.08, 0.26)	0.16 (0.09, 0.30)
Nugget (σ_{e}^2)	(0.17 (0.09,0.35)	0.15 (0.08, 0.30)	0.14 (0.08, 0.27)
Range, (in km)	5.04 (0.86, 49.55)	10.26 (0.86, 50.96)	3.90 (0.85, 50.43)

Figure 4.2 shows excess mortality (at different scales) that is attributed to malaria transmission intensity. Across all the age groups, there is a positive correlation between all-cause mortality and EIR. The highest burden was among the post-neonates followed by the neonates. This figure (4.2) indicates also that excess mortality curve becomes almost flat in all age groups as the number of infective bites go beyond 50 per person per month.

Presence of a net (ITN) in a household was associated with a reduction in all-cause mortality in the NHDSS for the three age groups. There was a positive trend in the protective effect of nets on children's age. The protective effects of nets on all-cause mortality were 15%, 21% and 50% for neonates, post-neonates and children between one and four years respectively.

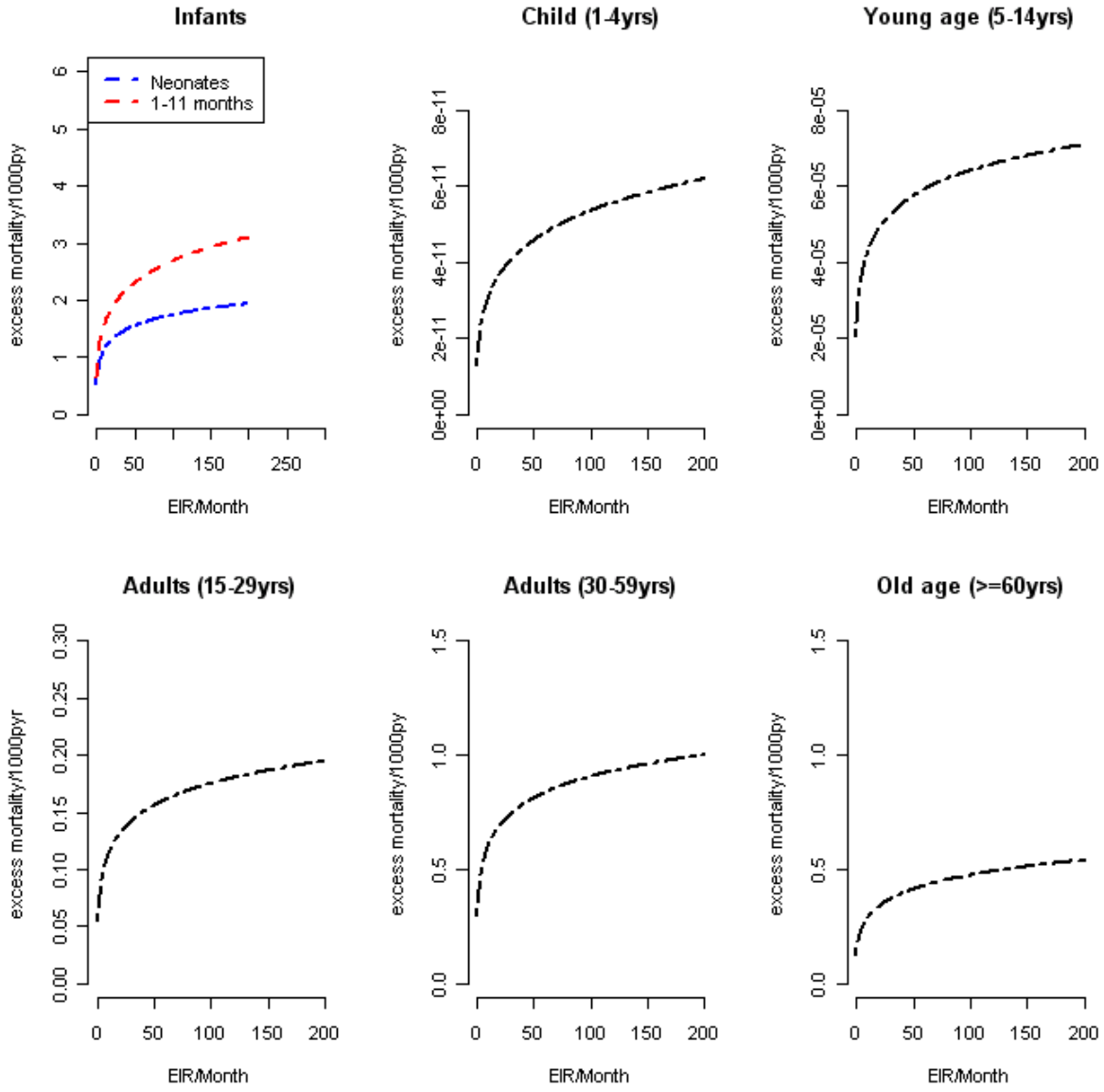


Figure 4. 2: Age specific malaria attributable mortality by EIR

4.4 Discussion

This study examined the relationship between mortality and malaria exposure in the NHDSS. Our analysis reports a decline in all-cause mortality rates across all the age groups during the 3 years of the MTIMBA project. The largest decline in mortality was observed among adults (15-29 years) followed by children between one and four years. A consistent decline in mortality

especially among the under fives has previously been reported in the district of Kassena-Nankana (Adjuik et al., 2010) and for the entire country. According to the Ghana Demographic and Health survey 2008, the under five mortality declined between 2003 and 2008 from 111 to 80 per 1000 live births respectively (GSS and GHS, 2009). The persistent reduction in mortality in this remote area may be partly associated with the research activities carried out by the NHRC (Appawu et al., 2004; Owusu-Agyei et al., 2007). The NHRC has tested low-cost interventions namely; vitamin A supplementation (Ghana VAST Study Team, 1993), permethrin-impregnated bed nets (Binka et al., 1996) and posting health workers in the communities that have been associated with reduced morbidity and mortality in the area (Binka et al., 2007; Nakamura et al., 2011).

As in other malaria endemic areas, mortality is highest among children under one year (Abdullah et al., 2007; Adjuik et al., 2006; Carneiro et al., 2010). (Figure 4.1) clearly show the highest rates among the infants (0-11 months) followed by the elderly population (60 years plus), a pattern similar to what has been observed from other sites with different transmission intensities (Adjuik et al., 2006; 2013; Rumisha, 2013). Although NHDSS experienced a decline in malaria transmission intensity during the MTIMBA project period, the third year EIR estimate was more than 100 infective bites per person (Kasasa et al., 2013) and the highest mortality burden remained in the same age group. Even in Western Kenya, where EIR was reduced to a single digit after promoting ITNs (Lindblade et al., 2004), malaria burden was still clustered in children under five years. This is similar to other SSA countries where malaria interventions are ongoing (Carneiro et al., 2010).

The Bayesian geostatistical spatio-temporal models showed a positive log-linear relationship between all-cause mortality and malaria exposure in all the age groups but the association was

only important among children (1-4 years) and old people (60 years plus). These findings confirm a positive relationship between mortality and transmission intensity among the infants that was derived from meta analyses (Ross and Smith, 2006; Smith et al., 2001). A positive log-linear relationship among the under fives was also observed in the 3 MTIMBA sites of Rufiji, Kisumu and Manhica (Rumisha 2013; Amek, 2013; Kasasa in preparation). The malaria risk estimates (hazard ratios) for the 4 sites are comparable.

The findings indicated an increasing trend in the magnitude of hazard ratios with age among the under-fives and a decline between 5 and 59 years, a pattern which is similar to what was observed in Manhica HDSS. The data from Rufiji showed a decline in magnitude of hazard ratios with age group among the children < 5 and individuals who were above 14 years. Analysis for the Kisumu data showed a consistent decline of hazard ratios among people aged 15 years and above.

Bayesian analysis identified children between 1 and 4 years with the highest hazard ratio of 13% associated with an e-fold increase in malaria transmission intensity. The same age category was reported in the Kisumu and Manhica analyses to have the highest mortality risk. The Rufiji HDSS which has transmission intensity comparable to the NHDSS one identified a different age group of 5 to 14 years with the largest hazard ratio. Like the data analyses in the other three MTIMBA sites, the NHDSS data identified the highest excess mortality burden among the infants. This implies that people in these areas still acquire their natural immunity at an earlier age despite the ongoing malaria control interventions that are targeting reducing transmission intensity.

Possession of a mosquito net in a household is associated with a reduction in all-cause mortality among the under fives in northern Ghana. The data show a protective effect of nets among all the

under five children similar to what was observed in Rufiji and Kisumu HDSS. The data used in this analysis were based on the presence of a net in a compound but not use. Although more than half of the compounds possessed atleast a mosquito net, this does not guarantee continuous use (Baume and Marin, 2007; GHS NMCP-Report, 2008; Githinji et al., 2010; Korenromp et al., 2003). Even in places with very high coverage, only a small proportion of the nets were treated and most nets had holes (Maxwell et al., 2006). Low use has been observed among children between 4 and 15 an issue that is likely to shift malaria burden among old children in highly endemic areas.

In conclusion this study reported a positive association between all-cause mortality and malaria transmission intensity in all age categories. Such relationships especially among the adults should be interpreted carefully. People in endemic areas develop clinical immunity at an early age and severe malaria cases usually reduce with increasing age. Further site-specific analysis followed by meta-analysis will provide proper evidence about mortality-malaria transmission relationship across different age groups.

4.5 Appendix 3

Age-specific spatio-temporal mortality model

Survival models are appropriate in analyzing mortality data and they always assume continuous follow-up time (Cox, 1972). Such models assume time-invariant predictors, yet malaria transmission intensity is heterogeneous. In the presence of time-dependant covariates, survival models are based estimated via logistic regression where discrete follow-up time is assumed (Allison, 1982; Manda and Meyer, 2005; Singer and Willet, 1993). The occurrence of each event is recorded sequentially as dummy variable at each observed time point. Monthly time intervals were created which allowed the use of EIR at that particular period.

The age-specific mortality data were therefore treated as Bernoulli outcome modelled separately via logistic regression. We assumed that status (dead/alive) Y_{ijt} of the child i status in village j at time t arises from Bernoulli distribution; that is $Y_{ijt} \sim Be(p_{it})$ with parameter p_{it}, the probability of dying. To account for spatial, temporal and non-spatial variation in the data, we introduced village, monthly time and non-spatial random effects ϕ_j, ε_t and e_j respectively. These were modelled together with covariates X_{it} (EIR and Age) on the logit scale as $\log it(p_{ijt}) = X_{it}^{T}\beta + \phi_j + \varepsilon_t + e_j$, where $\beta = (\beta_1, \beta_2, ..., \beta_k)^T$ is a vector of regression coefficients. The spatial random effects are assumed to originate from a Gaussian spatial process with zero mean and variance-covariance matrix Σ, where $\phi_i = (\phi_1, \phi_2, ..., \phi_n)^T \sim N(0, \Sigma)$. The covariance between any particular pair of compounds was assumed to be a function of distance between the locations, that is $\Sigma_{ij} = \sigma_\phi^2 \exp(-d_{ij}\rho)$ where d_{ij} is the Euclidean distance between locations i and j, σ_ϕ^2 is the spatial variance and ρ is the correlation decay

parameter with the range defined as $3/\rho$. The monthly time random effects were modelled by first order autoregressive process [AR (1)] with variance σ_{ε}^2 which allows correlation between consecutive time periods (Hay and Pettitt, 2001). Non-spatial random effects e_t with zero mean and variance σ_e^2 were included into the model to account for unexplained variation in the data. The EIR covariate was added to the model on a log scale via a normal distribution, that is $EIR \sim N(\log(EIR), SD))$; where $\log(EIR)$ and SD are the corresponding posterior mean and variance obtained at each compound.

We applied Bayesian inference by combining likelihood function and prior distributions to form the posterior distribution that was used in estimating model parameters. Prior distributions specified for all model parameters were; normal with zero mean and large variance for regression coefficients $[(\beta);\ \beta \sim N(0,10^3)]$ inverse gamma for the variances $[(\sigma_{\phi}^2, \sigma_{\varepsilon}^2, \sigma_e^2);\ \sigma^2 \sim IG(2.01,1.01)]$ variance σ_{ϕ}^2 and uniform for decay parameter $[\rho \sim U(a\ , b\ )]$. That is $\underline{\beta} \sim N(0,10^2)$, $\sigma_{\phi}^2 \sim IG(2.01,1.01)$ and $\rho \sim U(a,b)$ with hyper-parameters a and b the minimum plus the maximum values for ρ respectively. Models were implemented in OpenBUGS version 3.1.1 software where parameters were estimated using Markov Chain Monte Carol (MCMC) simulation algorithm. We used a two chains sampler with an initial burn-in of 10000 iterations.

Chapter 5: The effect of malaria transmission intensity on mortality in Manhiça Demographic Surveillance site, Southern Mozambique

Authors: Simon Kasasa[1, 2,3], Manhiça Team, Thomas Smith[1,2], Penelope Vounatsou[1,2*]

[1]Swiss Tropical and Public Health Institute, Socinstrasse 57, P.O. Box, 4002 Basel, Switzerland
[2]University of Basel, Petersplatz 1 P.O. Box 4003 Basel, Switzerland
[3] School of Public Health, Makerere University college of Health Sciences, P.O Box 7072, Kampala, Uganda

*Corresponding Author

This manuscript is prepared for submission to PLOS ONE Journal

Abstract

Reducing under-five mortality is one of the millennium development goals (MDG) Sub-Saharan Africa (SSA) is striving to achieve. Malaria is a major cause of mortality especially in younger age groups, but the quantitative relationship between mortality and malaria transmission intensity is unclear because of poor quality data in most endemic countries. The Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project collected malaria exposure data and linked this to the routine mortality surveillance of selected health and demographic surveillance sites (HDSS) to generate data to clear the mortality-malaria transmission relationship. In this sub-study we analyzed the entomological data from Manhiça HDSS to obtain malaria exposure surfaces which were linked to mortality data of all age groups. Bayesian geostatistical and temporal models were employed to predict entomological inoculation rate (EIR) at unsampled locations. Separate age-specific survival models were approximated via logistic regression incorporating the predicted EIR as a covariate with a measurement error. *Anopheles funestus* (85%) mosquitoes dominated the collections. The annual EIR declined from 43 to 3 infective bites with peaks in the wet months. Annual mortality rates were higher among infants and old ages above 60 yrs. The annual infant mortality rates (0-11 months) ranged between 90 and 104 per 1000 person-years (person-years) during the three years of the study. The annual child mortality rate declined from 23 to 19 per 1000 person-years. Model-based results indicated a positive log-linear relationship between all-cause mortality and malaria exposure across all age groups namely; neonates (hazard ratio, HR=1.11; [95% CI: 0.85, 1.4]), post-neonates (HR 1.13;[95% CI: 0.91, 1.36]), child (HR 1.25;[95% CI: 1.07, 1.44]), young (HR 1.22;[95% CI: 0.95, 1.55]), adults (HR 1.03;[95% CI: 0.93, 1.15]) and old people (HR 1.13;[95% CI: 0.99, 1.29]). Children (1-4 years) had the highest HR which was statistically important. A positive association between mortality and EIR among the under fives, is consistent with what was reported from the MTIMBA sites of Rufiji in Tanzania and Kisumu, Kenya. Adults' results need to be interpreted carefully because people in Manhiça are expected to acquire immunity at an early age due to the continuous exposure to *Plasmodium* infections.

5.1 Introduction

In Sub-Saharan Africa (SSA), malaria has been identified as one of the major causes of death especially among children under five years (Bardají et al., 2011; Black et al., 2010; Bryce et al., 2005). Human beings become infected with malaria parasites after being bitten by an infective anopheles mosquito. The disease is endemic in SSA. Some countries implementing interventions that are targeted towards achieving the United Nations Millennium Development Goals have reduced malaria morbidity and mortality (Okiro et al., 2009; Rajaratnam et al., 2010). Such changes have been partly attributed to improvement in health systems factors especially the scaling-up of effective malaria treatment and the use insect-treated nets (ITN) (Lengeler et al., 1998; Nyarango et al., 2006; Phillips-Howard et al., 2003; Steketee and Campbell, 2010). Although benefits from malaria programs are clear, the question on how mortality relates to transmission intensity has not been fully answered. This is partly due to lack of required data. Entomological data are expensive to collect and the mortality data are lacking due to inefficient vital registration systems found in many SSA countries. Early studies based on either meta-analysis (Ross and Smith, 2006; Smith et al., 2001; Snow and Marsh, 1995) or linking Demographic and Health Surveys (DHS) mortality and malaria risk from Malaria Risk in Africa (MARA) databases(Gemperli et al, 2004) provided inconsistent results about the malaria transmission-relationship. A clear understanding of mortality-malaria transmission intensity relationship would assist in implementing targeted interventions that will improve child survival. It will also help to establish whether the current malaria interventions have not caused a shift in age for acquiring natural immunity.

In early 2000, the INDEPTH network malaria working group integrated the Malaria Transmission Intensity and Mortality Burden Across Africa (MTIMBA) project in the routine

activities of selected health and demographic surveillance sites (HDSS) to generate reliable data. Entomological data were collected every two weeks over a large number of randomly selected households within each site. Mortality data were collected through routine demographic surveillance where verbal autopsy (VA) tools were applied to determine the most probable cause of death (Adjuik et al., 2006; Joshi et al., 2009).

Analyses of the Rufiji in Tanzania (Rumisha, 2013) and Kisumu in Kenya (Amek, 2013). MTIMBA data showed a positive log-linear relationship between all-cause mortality and malaria transmission intensity among the children under five years. The highest risk of dying from the disease after malaria exposure was found among children (1-4years) and school children (4-15 years) for Kisumu and Rufiji sites respectively. Although these two sites have different malaria transmission intensities and mortality risks, both experience temporal and seasonal variations. Further analysis of site-specific MTIMBA data using similar methods will help to clarify the mortality-transmission relationship.

This paper reports estimates of EIR from Bayesian geostatistical models, fitted separately for sporozoite rate (assumed to be binomially distributed) and mosquito density data (negative binomially distributed) and further established its relationship to age specific mortality using data collected from Manhiça HDSS.

5.2 Methods

5.2.1 Setting

Manhiça HDSS located in Manhiça district, Southern Mozambique was established in 1996 and currently covers 500 square kilometres. The area is divided into two ecological zones, the fertile lowlands and an escarpment where most people reside. The lower floodplains favour mosquito breeding all the year. There are predominantly two climatic seasons, hot and wet between

October to April plus a cool and dry windy period for the rest of the year. The region experiences perennial malaria transmission with a seasonal variation. *Anopheles funestus* is the main mosquito species in this area. The estimated EIR for 2002 was 15 infective bites per person per year (ib/p/y) (Aranda et al., 2005). The 2006, neonatal and infant mortality rates were 26 and 75 per 1000 live birth respectively (Menéndez et al., 2010). Data from VA interviews indicate that malaria accounts for almost a quarter of all paediatric deaths in this area (Sacarlal et al., 2009). The district is served by two referral hospitals (The Manhiça district hospital and Xinavane Rural hospital) and 10 other peripheral health facilities**.** Further details of the district and the HDSS have been documented elsewhere (Saúte et al., 2003).

5.2.2 Data types and sources

i) Entomological data

Mosquito collection was performed over a three year period (October 2001 and September 2004) following the MTIMBA protocol for monitoring transmission data. The Centre for Disease Control (CDC) light trap catches were set up in randomly selected households within identified clusters. Each trap was positioned indoor at about 1.5m above the floor next to the bed of an "indexed" person from sundown to sunrise. All mosquitoes were transported from the field to the laboratory for identification and testing. Heads and thoraces of light trapped anopheles were checked for the presence of circumsporozoite antigen of *Plasmodium falciparum* using enzyme linked immunosorbent assay (ELISA). EIR was generated as a product of sporozoite rate and man-biting rate. Due to the correlation between the number of mosquitoes captured using light trap catches and human biting catches, the later were estimated by including a conversion factor (Lines et al., 1991). These estimates were further divided by the number of light traps to obtain the man-biting rates.

ii) Climate and environmental data

The climatic and environmental covariates used were day and night land surface temperature (LST), rainfall, enhanced vegetation index (EVI), normalized difference vegetation index (NDVI), altitude and distance to water bodies. The latter was extracted from locally generated maps by the HDSS. The sources and the spatio-temporal resolutions of the data for the remaining variables have already been documented in our previous work (Kasasa Simon, et al., in preparation). The climatic and environmental variables were processed at the locations where entomological data were available. For each location, temperature, rainfall and vegetation data were summarized by month and year during the duration of MTIMBA project.

iii) Mortality data

Mortality data are routinely collected by the HDSS using standard methods and were extracted from the HDSS database. Cause-specific mortality data were obtained using a modified verbal Autopsy (VA) tool derived from both, the World Health Organization (WHO) and the HDSS standard VA questionnaires. Trained fieldworkers translated the questionnaires from Portuguese into local language (Xangana) in order to capture all the information surrounding the death. The VA process in the HDSS has been described elsewhere (Sacarlal et al., 2009). Geo-reference compounds where mortality was monitored during the project period were linked to the nearest predicted EIR using ArcGIS software.

5.2.3 Data analysis

i) Sporozoite and mosquito density data

Sporozoite rate data were modelled in our previous work via logistics regression (Kasasa et al. 2012). Non-spatial negative binomial regression models were used to analyse mosquito density data. Zero inflated models were fitted to account for the large number of locations with no

mosquitoes (55%; *An.funestus* and 84%; *An.gambiae*). Exploratory analyses carried out in STATA software assessed the effect of elapsing time (lags) between climatic suitability and malaria transmission using Akaike's information criterion (AIC)(Hoeting et al., 2006). A total of five summary estimates were computed for each of the environment factors based on mosquito collection month in a year (current month of collection; previous month; previous two months; average of the current and previous month; average of the current and previous two months). Temperatures were measured by LST day, night and average. Seasonality was taken into accounted by either a binary variable (wet/dry) or trigonometric functions with (i) one cycle indicating a single transmission season or (ii) two cycles corresponding to two transmission seasons per year. A suitable combination of environmental and climatic predictors of mosquito density by vector species was identified using AIC.

Bayesian geostatistical formulations of the above models were fitted to take into account spatio-temporal correlation. In each model, compound-specific random effects were included. They were assumed to be latent observations from a multivariate Gaussian spatial process with zero mean. The covariance of the process assumed an exponential correlation function of distance between any pair of compound locations. First-order autoregressive terms were included to model temporal correlation. Any remaining non-spatial variation (nugget parameter) was considered by additional location random effects which are independent and normally distributed with zero mean. Random effects and the covariates were modelled on a log scale of the mean of the negative binomial distribution. Bayesian kriging was applied to predict mosquito density over a grid of 2100 pixels with 250m by 250m spatial resolution. The analysis was carried out for each mosquito species (i.e. *An.funestus* and *An.gambiae*) separately.

Parameters were estimated using Markov Chain Monte Carol (MCMC) simulation algorithm (Gelfand and Smith, 1990). Samples from the posterior predictive distributions of SR and mosquito density were multiplied at each location to generate spatially explicit surfaces of EIR for each species. A conversion factor of 1.605 was used to convert densities to man-biting rates (Lines et al., 1991). Monthly maps for the total EIR were obtained using ArcGIS software. Mathematical description for all models and implementation details are similar to those described in our previous work (Kasasa et al., 2013) **.**

ii) Mortality data

These data were categorized according to age namely; neonates (0-28 days), post-neonates (1-11 months), children (1-4 years), young people (5-14 years), adults (15-59 years) and old age (60 years and above) partly due to the variations in mortality and its' determinants among different age groups (Becher et al., 2008). Crude estimates and modelling were conducted according to the predefined age categories. Mortality rates (MR) were computed by taking the total number of deaths divided by the total person time at risk. The annual rates were expressed per 1000 person-years (person-years). Non-spatial Cox regression models were initially fitted to all the age groups for exploratory analysis Survival models were approximated via logistic regression which assumes a discrete time to death(Manda and Meyer, 2005; Singer and Willet, 1993). Bayesian geostatistical models were fitted to the data in order to control for space and time correlation. In each model, village-specific, monthly time and non-spatial (nugget) random effects were included to account for spatial, temporal and random (nonspatial) variations in mortality data. All the corresponding random effects were defined in similar way as described in the zero-inflated negative binomial Bayesian geostatistical model for mosquito density above. These random effects and covariates (age and EIR) were modelled on a logit scale for mortality

data. Mortality was related to logarithmic transformed EIR of the previous month. Predicted EIR was incorporated in the model as a covariate with measurement error to account for the prediction uncertainty (Gemperli, 2003). Mathematical description for all models and implementation details are indicated in the appendix 4.

ii) Excess mortality

Model coefficients from the final survival model were used over a range of EIR between zero to 300 infective bites per person per month to estimate excess mortality. The excess mortality rate was computed as the difference between mortality rate when the transmission intensity (EIR) is more than zero and at zero EIR. The risk of death was converted into rates and a Taylor series approximation was used to generate mortality at zero EIR. All the details of calculating excess mortality have been documented in the analyses of Kisumu and Rufiji HDSS data (Amek, 2013; Rumisha, 2013)

5.3 Results

5.3.1 Description of density data

From 2918 unique compounds, a total of 18923 mosquitoes were caught with An. *funestus (85%)* dominating the traps and the rest were *An. gambiae (15%).* The mean biting rates per person per night from first to third year were 5, 3 and 2 respectively. During the survey period, the mean biting rates per month varied with seasonal changes. Most bites were observed during the wet and warm period.

5.3.2 Description of EIR data

The crude annual entomological inoculation rates from year one to three were 43, 28 and 3 infective bites per person respectively for the combined vector species. The corresponding annual EIR estimates for *An. funestus* were, *37, 25 and* 3 infective bites per person, while *An.*

gambiae contributed 6, 3 and zero infective bites respectively. For the two vector species, year one had the highest number of bites. The monthly distribution of EIR and the rainfall patterns in Figure 5.1 show a decline in EIR over time. In year 3, most of the mosquitoes caught were not infectious as the monthly EIR ranged between zero and one infective bite per person.

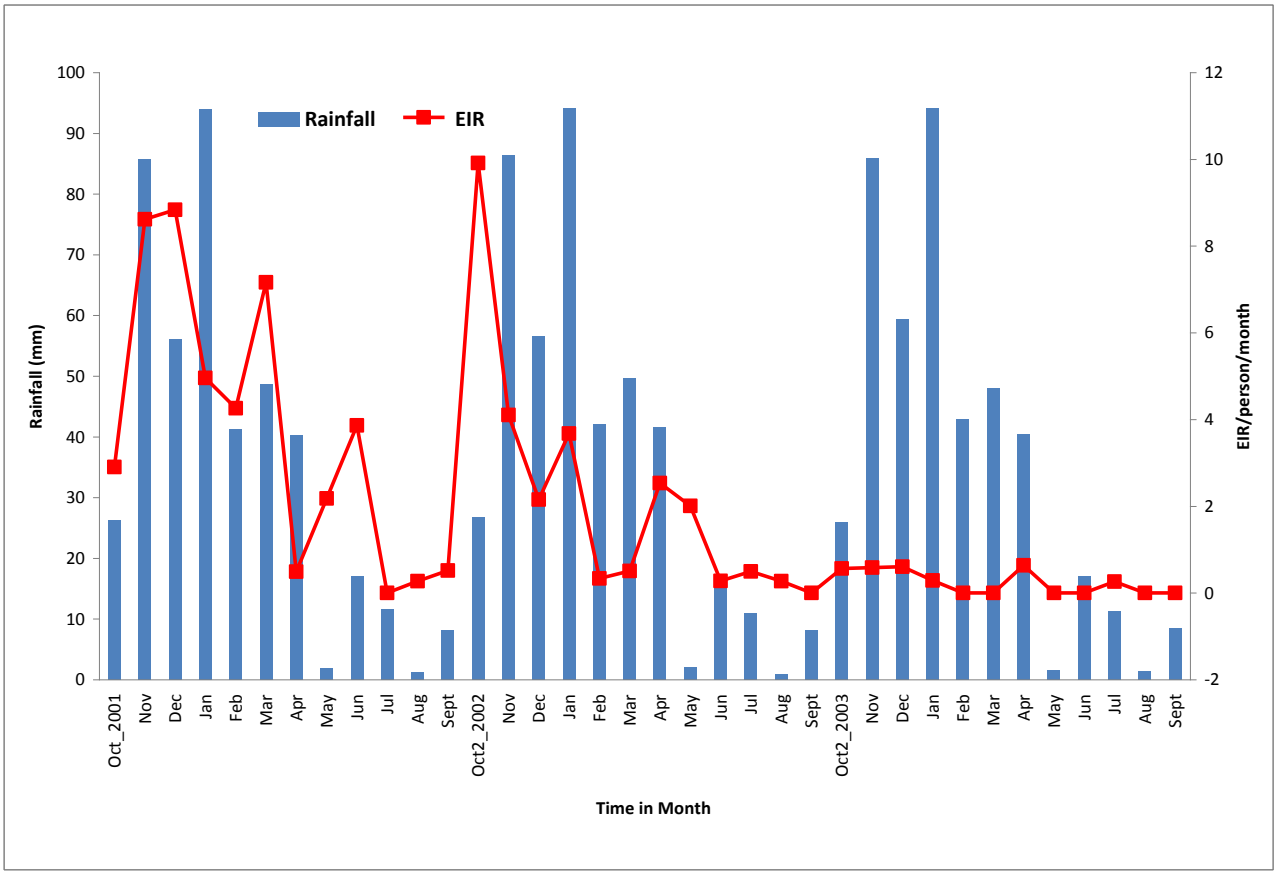


Figure 5. 1: Monthly rainfall and infective mosqiote bites

5.3.3 Description of mortality data

A total of 45032 individuals from 7835 geo-referenced compounds were followed-up and 2027 deaths with 99598.43 prys were registered during the survey period. There were more females (54%) in the HDSS. However, deaths were evenly distributed between gender (50.5% male). The crude mortality rates by age categories are summarized in table 5.1.

Table 5. 1: Crude child all-cause mortality rates by year (2001/02- 2003/04)

	Annual mortality rate per 1000 person-years by age groups									
Year	0-11 months		*1-4 years*		*5-14 years*		*15-59 years*		*60 yrs plus*	
	Deaths (pyrs)	MR	Deaths (pyrs)	MR	Deaths (pyrs)	MR	Deaths (pyrs)	MR	Deaths (pyrs)	MR
1	108 (1096.6)	98.5	99 (4303.9)	23.0	37 (8202.1)	4.5	299 (15991.3)	18.7	141 (2386.9)	50.1
2	128 (1228.7)	104.2	94 (4233.5)	22.2	21 (8539.3)	2.5	286 (16273.9)	17.6	155 (2413.3)	64.2
3	113 (1244.2)	90.8	78 (4158.3)	18.8	29 (8707.7)	3.3	312 (16438.4)	19.0	127 (2395.3)	53.0
Total	349		271		87		897		423	

pyrs: person-years at risk; MR: all-cause mortality rate

The risk of dying was higher among children who had not attained their first birthday. IMR varied between 91 and 104 per 1000 person-years during the MTIMBA period. It increased between year one and two but later declined by 13% between second and third years. The old people in this district also experienced a higher mortality rate of 50, 64 and 53 per 1000 person-years for the first to last year respectively. This is second age group in the area with high mortality rates. For the three years, young people (5-14 years) experienced the lowest mortality risk from the first to third year of 5, 3 and 3 per 1000 person-years respectively. Annual child mortality rate decreased consistently from 23 to 19 per 1000 person-years.

A total of 1732 (85.4%) VA interviews were conducted to identify the likely cause of death. Malaria was assigned to 198 deaths (11.4%). Other diseases were not defined in the database. No death under five years was assigned to malaria. Majority of malaria deaths (68.7%) were between 15 and 59 years old.

5.3.4 Model-based results: Mosquito density data

Non-spatial models identified the following significant predictors of mosquito density of A*n. funestus* namely; previous month EVI, day and night temperatures plus rainfall over the past two months prior the project. Similarly A*n. gambiae* mosquito density were driven by current EVI, night LST plus rainfall and day temperature for the past two months prior the survey. Posterior estimates from geostatistical zero-inflated negative binomial models are summarized by mosquito species in table 5.2.

Table 5. 2: Multivariate space and time analysis for mosquito density by species

Parameters	*An. funestus*		*An. gambiae*	
	Co-efficients		Co-efficients	
	Median	95% CI	Median	95% CI
Intercept	1.80	(-0.01, 4.21)	-1.08	(-4.73, 1.21)
Altitude	0.003	(-0.004, 0.01)	0.005	(-0.01, 0.02)
Distance to water bodies	-1.06	(-1.55, -0.18)	-1.26	(-1.83, -0.84)
NDVI/EVI	1.30	(-0.01, 2.71)	1.75	(0.03, 4.47)
Rainfall	-0.001	(-0.01, 004)	0.01	(0.001, 0.02)
Season(Wet)	0.04	(-0.55, 0.94)	0.88	(-0.28 , 1.85)
Day temperature	0.03	(-0.03, 0.09)	0.11	(0.01, 0.20)
Night temperature	-0.04	(-0.15, 0.05)	-0.20	(-0.38, -0.06)
Survey period	-	-	-	-
2	-0.41	(-1.50, 1.43)	-0.81	(-1.59, 1.39)
3	-0.75	(-1.89, 1.38)	-1.44	(-3.23, 1.52)
Variances				
Spatial (σ_{ϕ}^2)	2.16	(1.31, 4.08)	1.68	(0.98, 3.53)
Temporal (σ_{ε}^2)	0.60	(0.34, 1.17)	1.09	(0.60, 2.09)
Range (in km)	8.05	(4.68, 15.44)	4.27	(2.11, 10.50)
Dispersion parameter (r)	0.47	(0.43, 0.51)	0.30	(0.26, 0.36)

We only presented results from mosquito density models without a nugget because of they fitted the data best (other results not shown). For *An. funestus* data, distance to water bodies was the only significant covariate in the model. The model suggests that *EVI*, elevation, climatic season

and day temperatures were positively associated with density. A negative association was observed with shortest distance to water bodies, rainfall, night temperatures and survey time period. The decay parameter (ρ) had a posterior median of 41.7 (95% Cl: 21.8, 71.8) which corresponds to a distance of 8.1 km (95% CI: 4.7km, 15.4km) at which the spatial correlation is less than 5%. The spatial variation (σ_{ϕ}^2 =2.2, [95% CI: 1.3, 4.1]) was larger than temporal one (σ_{ε}^2 =0.6, [95% CI: 0.3, 1.2]).

The analysis of the *An. gambiae* data showed that, EVI, altitude, rainfall, season and day LST covariates were positively related to density. Similarly, distance to water bodies and night temperature were negatively associated to mosquito density. The most important environment predictors of the mosquito density were distance to water bodies, EVI, rainfall plus day and night LST. Spatial variance (σ_{ϕ}^2 =1.7, [95% CI: 1.0, 3.5]) was larger than temporal one (σ_{ε}^2 =1.1, [95% CI: 0.6, 2.1]). The minimum distance at which the spatial correlation was below 5% was 4.3km (95% CI: 2.1km, 10.5km) for both species. Over-dispersion was present for both A*n. funestus* (r=0.5, [95% CI: 0.4, 0.5]) and An. *gambiae* (r=0.3, [95% CI; 0.3, 0.4]) mosquito density data.

5.3.5 Model-based results: EIR estimates

The maps of predicted monthly EIR for the first year are given in Figure 5.2. Maps clearly show a seasonal and spatial pattern, ranging from almost no infective bites in dry and cool months to the highest number of infective bites in the wet and warm months of October to April. For the entire project period, the predicted EIR for the month of July was almost zero. Areas in the northern and eastern part of the Manhiça receive slightly high number of infective bites compared to the rest of the HDSS. Our models were able to predict high transmission in areas close to the Incomati River.

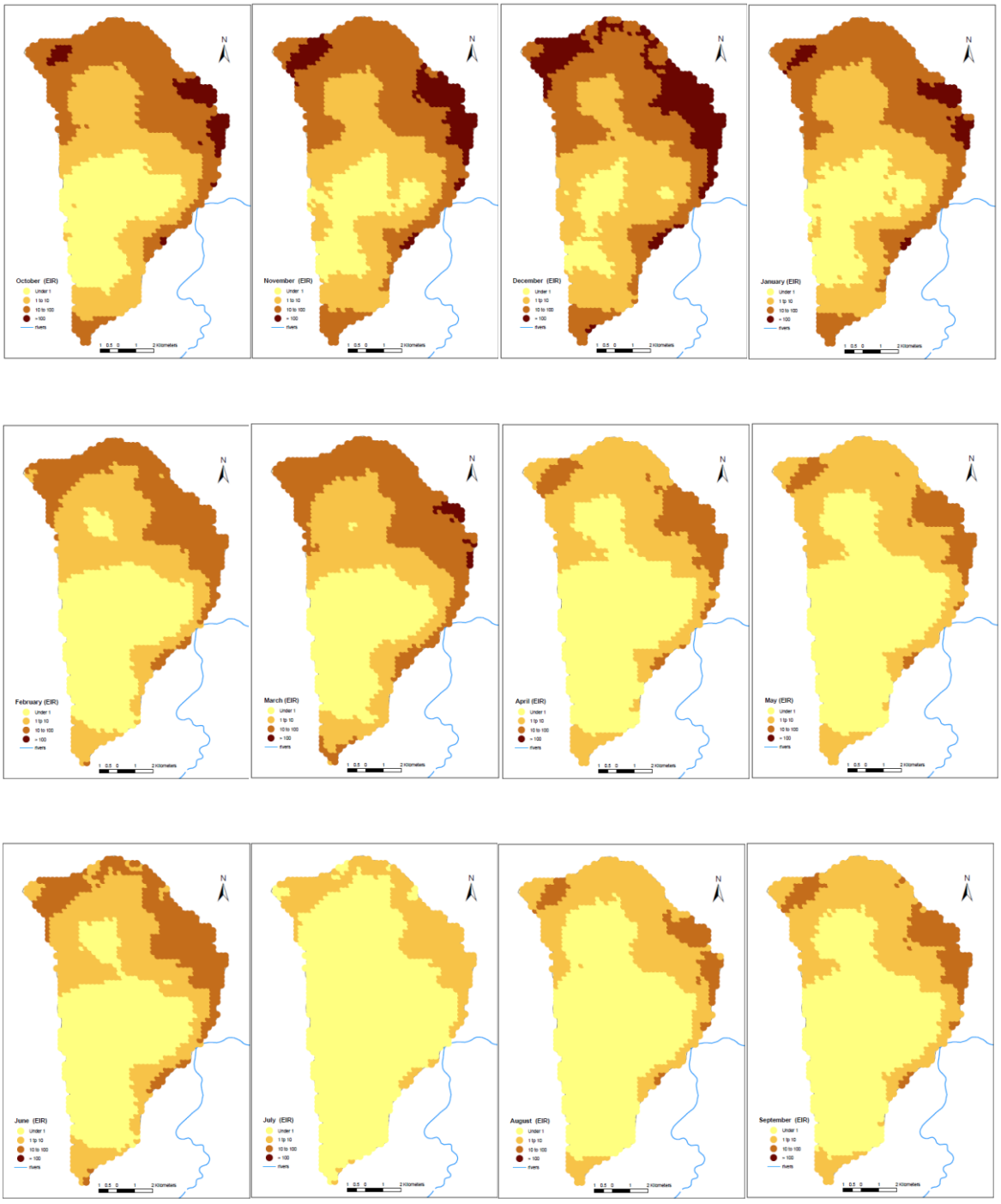


Figure 5. 2: Predicted monthly entomological inoculation rate for the first year

5.3.6 Model-based results: mortality data

Table 5. 3: Spatio-temporal multivariate posterior estimates for all-cause mortality

Parameters	*0-28days*	*1-11months*	*1-4 yrs*	*(5 - 14 yrs)*	*(15-59 yrs)*	*60yrs plus*
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Constant	0.17 (0.09, 0.29)	0.01 (0.003, 0.01)	0.01 (0.003, 0.02)	0.0003 (0.0002, 0.001)	0.0002 (0.0001, 0.002)	-
Log EIR*	1.11 (0.85, 1.39)	1.13 (0.91, 1.36)	1.25 (1.07, 1.44)	1.22 (0.95, 1.55)	1.03 (0.93, 1.15)	1.13 (0.99, 1.29)
Age**	0.82 (0.80, 0.84)	0.90 (0.86, 0.94)	0.95 0.94, 0.96)	0.92 (0.86, 0.98)	1.05 (1.04, 1.06)	1.07 (1.06, 1.08)
Variances						
Spatial (σ^2_ϕ)	0.25 (0.12, 0.63)	0.22 (0.11, 0.51)	0.26 (0.12, 0.63)	0.32 (0.14, 0.83)	0.17 (0.09, 0.36)	0.20 (0.01, 0.46)
Temporal (σ^2_ε)	0.22 (0.12, 0.46)	0.23 (0.12, 0.45)	0.19 (0.11, 0.38)	0.19 (0.10, 0.40)	0.11 (0.07, 0.19)	0.13 (0.08,0.25)
Nugget (σ^2_e)	0.23 (0.11, 0.54)	0.21 (0.11, 0.21)	0.23 (0.11, 0.53)	0.33 (0.14, 0.84)	0.15 (0.08,0.29)	0.19 (0.10, 0.41)
Range (in km)	1.20 (0.44, 10.67)	1.03 (0.44, 10.56)	0.78 (0.44, 7.8)	1.41 (0.45, 10.6)	2.22 (0.45, 11.54)	1.12 (0.44, 10.8)

yrs -years ; * Natural logarithim of EIR; ** Units of age were based on model category (days, months and yrs)

Spatial-temporal modelling of the relationship between mortality included age and EIR covariates as shown in Table 5.3. The data shows a positive log-linear relationship between all-cause mortality and malaria exposure across the six age categories namely; neonates (hazard ratio; HR=1.11; [95% CI: 0.85, 1.4]), post-neonates (HR 1.13;[95% CI: 0.91, 1.36]), child (HR 1.25;[95% CI: 1.07, 1.44]), young (HR 1.22;[95% CI: 0.95, 1.55]), adults (HR 1.03;[95% CI: 0.93, 1.15]) and old people (HR 1.13;[95% CI: 0.99, 1.29]). The hazard ratios clearly refer to the effect of an e-fold change in EIR.

Figure 5.3 where excess mortality due to malaria exposure was estimated clearly show the same positive pattern across all age groups. This implies that as the number infectious bite per person per month increase, all-cause mortality also increases. The highest burden was among the post-neonates, neonates and young age (5-14 years). However, in all age groups, excess mortality tends to almost constant when EIR reaches 100 infective bites per person per month.

The data shows that the highest risk of dying associated with increased malaria exposure was among children between 1 and 4 years (25%) and least for the adults (3%).. The mortality effect due to increased EIR was similar for both the infants (1-11months) and old age people (60 years plus). The risk of dying from EIR exposure increased with age among the under fives. Age was significantly related to mortality in all the six groups.

The parameters σ_{ϕ}^{2} and ρ measure the spatial variance and rate of correlation decay. Our data indicates that the minimum distance at which the spatial correlation is less than 5% for all the age categories is between 0.8-2 km suggesting a weak geographical dependency. The variances of spatial heterogniety were 0.25(95% CI: 0.12, 0.63), 0.22(95% CI: 0.11, 0.51), 0.26(95% CI: 0.12, 0.63), 0.32(95% CI: 0.14, 0.83) , 0.17(95% CI: 0.09, 0.36) and 0.2 (95% CI: 0.01, 0.46) for neonates, infants, child young, adult and old people respectivley.Apart from one age category (1-4), the spatial variances were larger than the temporal for all the six data categories.

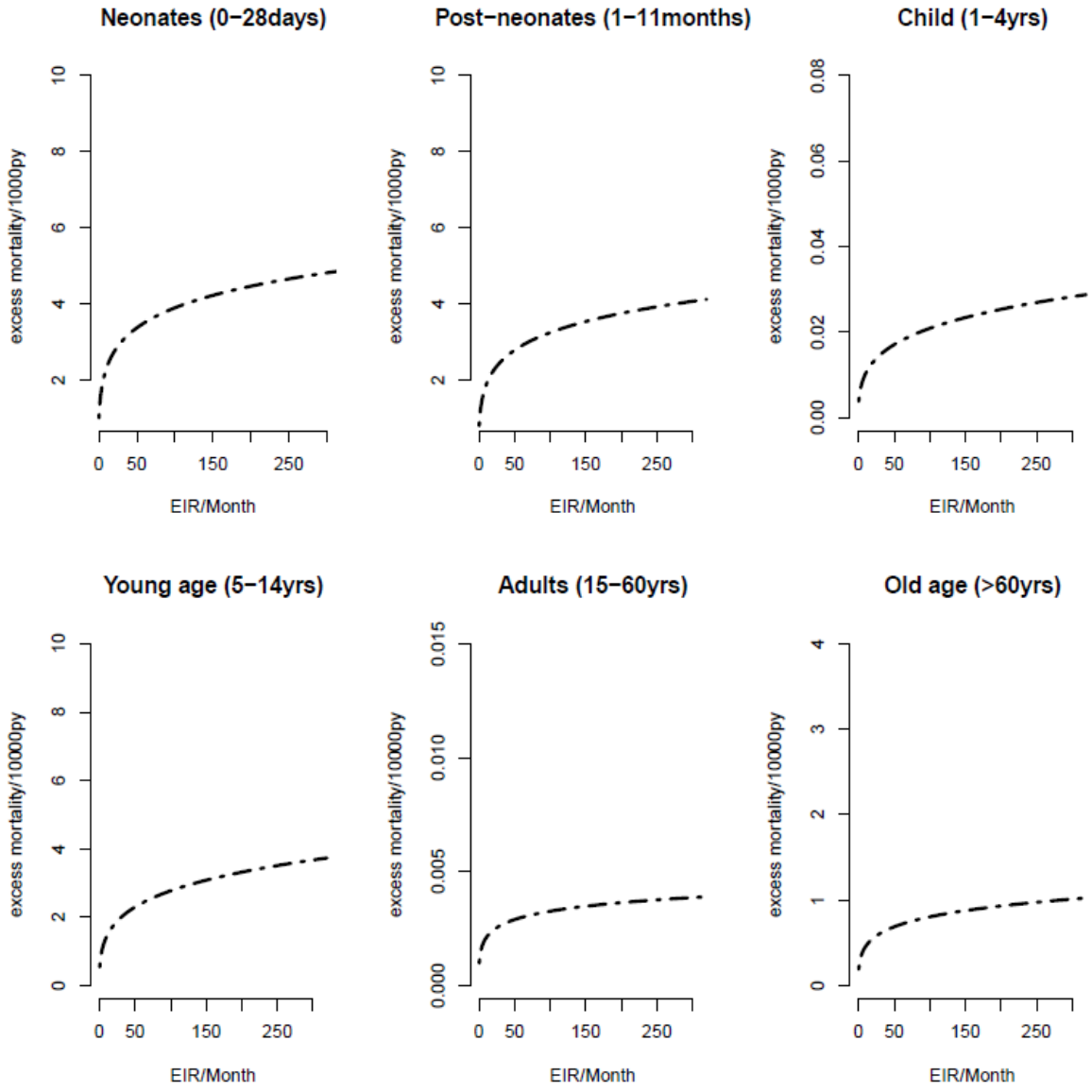


Figure 5. 3: Patterns of age specific excess mortality by transmission levels

5.4 Discussion

The present study assessed spatio-temporal variation in mortality and its relationship with malaria transmission intensity in southern Mozambique. An EIR exposure surface was estimated by fitting Bayesian geostatistical spatio-temporal models on sporozoite rate (logistic regression) and mosquito density (negative binomial regression) data. Although, the spatial pattern is similar

over time, transmission intensity is higher during the wet and warm months of October to April, implying that changes in climatic season do not modify the geographical distribution. Malaria transmission was higher in the north and north-eastern part of the HDSS. The EIR spatial pattern is similar to the spatial distributions of malaria incidence and mortality already observed in children living in the same area (Abellana et al., 2008; Escaramís et al., 2011). Spatial heterogeneity of malaria transmission intensity across Africa has been described (Hay et al., 2000) and confirmed in the MTIMBA sites of Rufiji (Rumisha et al. 2013) and Kisumu (Amek et al., 2012).

Although transmission in this region is moderate (Aranda et al., 2005), there was an observed consistent decline in intensity over the three years. A similar temporal trend was also reported in northern Ghana (Kasasa Simon, et al., 2013) but not in the two East African MTIMBA sites of Rufiji and Kisumu (Amek, 2013; Rumisha 2013). The decline could be attributed to control interventions especially the sporadic indoor residual spraying (IRS) that was introduced in Mozambique in the 1940s (Mabaso et al., 2004; Munguambe et al., 2011).

Seasonality in the MTIMBA data was modelled using either trigonometric sine/cosine functions (Rufiji and Kisumu) or an indicator for a dry/wet month (Navrongo) (Kasasa Simon, et al., 2013) and was present in all the three sites. In Mozambique, seasonality has already been reported in malaria incidence(Abellana et al., 2008) and maternal mortality (Romagosa et al., 2007) which is linked to the disease. The entomological data from the four MTIMBA sites had large number locations with zeros especially in the dry months. Zero-inflated formulations of geostatistical model are needed to estimate more accurately the frequency of excess zeros. The density data shows spatial variance larger than the temporal one for both species. Over 60% of the total

variation in the data was explained by space and a spatial correlation was even still strong at large distance.

According to the MTIMBA data, transmission in this district is mainly driven by A*n. funestus* malaria vector as previously reported (Aranda et al., 2005; Kloke et al., 2011). Mosquito abundance was mainly influenced by the nearest distance to water bodies such as the areas close to the Incomati River and the flood plains which act as breeding sites for mosquitoes throughout the year.

The vegetation index was positively associated with mosquito abundance for the two species. Vegetation is known to be directly related to other climatic factors such as rainfall. This is supported by the increase of EIR during the wet season. High mosquito densities and infective bites have been previously observed in the same region during the warm months of November to May (Aranda et al., 2005; Mendis et al., 2000).

During the MTIMBA project period, Manhiça HDSS recorded an increase in both infant and old age mortality between the first and second year which was followed by a drop in the final year. Child mortality (1-4 years) declined throughout the study period. The observed decline especially in infant mortality in southern Mozambique where ITN ownership is low (Chase et al., 2009) could be attributed to other malaria interventions that target pregnant women and their babies (Menéndez et al., 2010; Munguambe et al., 2011) and possibly improvement in health services. Like the infants, old people (60 years plus) in the HDSS experienced higher mortality rates during the study period compared to others. The age mortality pattern for Manhiça district where mortality declines after 4 years and rises up gradually from 15 years (Table 5.2) is similar to the one reported for Rufiji (Rumisha, 2013) ,Kisumu (Amek, 2013) and other HDSS in the

region (Adjuik et al., 2006). The increase in adult mortality in SSA is partly attributed to HIV (Blacker, 2004) and the VA confirmed this in Nouna HDSS (Becher et al., 2008).

Geostatistical and temporal survival models fitted on mortality data indicated that transmission intensity (log EIR) was positively associated with mortality in all the six age categories. The association especially among individuals >5 years, is not expected in higher transmission areas because of acquired immunity. A positive relationship has been also reported in previous infant mortality studies(Ross and Smith, 2006; Smith et al., 2001). Analysis from Rufiji HDSS reported a positive log-linear relationship with the under fifteen children only while Kisumu HDSS data showed the same effect but among the under fives. An inverse relationship was observed among the rest of the age groups in the two sites.

Our data show the highest increase in hazard ratio due to increase in malaria exposure was among children between 1 and four years (25%) followed by young people (22%) and the least in adults (3%). Amek et al, (Amek, 2013) identified the same age group with highest hazards of death. Although Kisumu HDSS observed larger effects of malaria exposure on mortality among children under five years than Manhiça, the transmission intensities in the two site sites are comparable (Amek et al., 2012). Data from another MTIMBA site of Rufiji reported that children 5-14 years had the highest hazard ratio due to increased malaria exposure.

The analysis indicated an increasing trend in the magnitude of hazard ratios with age among the under-fives and a decline between 5 and 59 years. The Kisumu data did not show a clear pattern among children but rather a consistent decline of hazard rate in individuals of 15 years old and above. The results from Rufiji indicated a decreasing trend in the magnitude of hazard ratios among the under-fives and individuals aged atleast 15 years.

Unlike the other two MTIMBA sites, in this analysis, we did not adjust for factors that have been associated with mortality namely, social economic status (Houweling et al., 2006), ITN ownership and use (Lengeler, 2004) because those data were not available. However our findings did not completely divert from those of Rufiji and Kisumu HDSS. We are currently analyzing MTIMBA data from Navrongo (Ghana) and Nouna (Burkina Faso) HDSS to get a better understanding of the mortality-transmission intensity relationship across different age groups.

5.5 Appendix 4

Spatio-temporal modelling of mortality

Survival models are appropriate in analyzing mortality data and they always assume continuous follow-up time (Cox, 1972). Such models assume time-invariant predictors, yet malaria transmission intensity is heterogeneous. In the presence of time-dependant covariates, survival models are based estimated via logistic regression where discrete follow-up time is assumed (Allison, 1982; Manda and Meyer, 2005; Singer and Willet, 1993). The occurrence of each event is recorded sequentially as dummy variable at each observed time point. Monthly time intervals were created which allowed the use of EIR at that particular period.

The age-specific mortality data were therefore treated as Bernoulli outcome modelled separately via logistic regression. We assumed that status (dead/alive) Y_{ijt} of the child i status in village j at time t arises from Bernoulli distribution; that is $Y_{ijt} \sim Be(p_{it})$ with parameter p_{it}, the probability of dying. To account for spatial, temporal and non-spatial variation in the data, we introduced village, monthly time and non-spatial random effects ϕ_j, ε_t and e_j respectively. These were modelled together with covariates X_{it} (EIR and Age) on the logit scale as $\log it(p_{ijt}) = X_{it}^{T}\beta + \phi_j + \varepsilon_t + e_j$, where $\beta = (\beta_1, \beta_2, ..., \beta_k)^T$ is a vector of regression coefficients. The spatial random effects are assumed to originate from a Gaussian spatial process with zero mean and variance-covariance matrix Σ, where $\phi_i = (\phi_1, \phi_2, ..., \phi_n)^T \sim N(0, \Sigma)$. The covariance between any particular pair of compounds was assumed to be a function of distance between the locations, that is $\Sigma_{ij} = \sigma_\phi^2 \exp(-d_{ij}\rho)$ where d_{ij} is the Euclidean distance between locations i and j, σ_ϕ^2 is the spatial variance and ρ is the correlation decay

parameter with the range defined as $3/\rho$. The monthly time random effects were modelled by first order autoregressive process [AR (1)] with variance σ_{ε}^2 which allows correlation between consecutive time periods (Hay and Pettitt, 2001b). Non-spatial random effects e_t with zero mean and variance σ_e^2 were included into the model to account for unexplained variation in the data. The EIR covariate was added to the model on a log scale via a normal distribution, that is $EIR \sim N(\log(EIR), SD))$; where $\log(EIR)$ and SD are the corresponding posterior mean and variance obtained at each compound.

Model fit and implementation

We applied Bayesian inference by combining likelihood function and prior distributions to form the posterior distribution that was used in estimating model parameters. Prior distributions specified for all model parameters were; normal with zero mean and large variance for regression coefficients $[(\beta);\ \beta \sim N(0,10^3)]$, inverse gamma for the variances $[(\sigma_{\phi}^2, \sigma_{\varepsilon}^2, \sigma_e^2);\ \sigma^2 \sim IG(2.01, 1.01)]$ variance σ_{ϕ}^2 and uniform for decay parameter $[\rho \sim U(a\ , b\ )]$. That is $\underline{\beta} \sim N(0,10^2)$, $\sigma_{\phi}^2 \sim IG(2.01,1.01)$ and $\rho \sim U(a,b)$ with hyper-parameters a and b the minimum plus the maximum values for ρ respectively. Models were implemented in OpenBUGS version 3.1.1 software where parameters were estimated using Markov Chain Monte Carol (MCMC) simulation algorithm. We used a two chains sampler with an initial burn-in of 5000 iterations.

Chapter 6: General discussion and conclusions

6.1 Discussion

This work contributes to the existing literature of estimating mortality related burden of malaria across age groups using longitudinal geostatistical data contributed by two sites namely; the Navrongo HDSS in Ghana and Manhiça HDSS in Mozambique. Specific contributions were made to (i) the descriptions of the HDSS data characteristics and relevant methods for analysis; (ii) spatially explicit estimates of malaria transmission intensity at monthly intervals; and (iii) the relationship between all-cause mortality and malaria transmission intensity across all age categories. High spatio-temporal resolution entomological inoculation rate (EIR) estimates were generated from sporozoite rate and mosquito density models that adjusted for climatic and environmental factors. During the assessment of all-cause mortality-malaria transmission intensity relationship, the data were aligned in space and time because entomological data were only captured in selected households while mortality was monitored in the entire HDSS over the MTIMBA project period. The estimated EIRs were included in age-specific mortality models with their respective measurement errors to adjust for prediction uncertainty.

This thesis contributes further to data-driven rigorous statistical methods for analyzing large entomological and mortality data that were collected while not aligned in both time and space (sparse and non-Gaussian data). The statistical methodologies applied generated information that is useful in the general understanding of how malaria transmission intensity influences mortality in sub-Saharan Africa where the disease is endemic. The work which was undertaken generated four manuscripts that constitute the main chapters (2 to 5) of the thesis. The first manuscript (chapter 2) discussed the characteristics of the MTIMBA project data, statistical issues and

proposed appropriate statistical methods (prepared for submission in Malaria). The third chapter applied geostatistical models to generated smooth monthly maps of EIR for Navorongo HDSS (published in *Malaria Journal*). Chapter 4 utilized predicted EIR surfaces to assess the mortality-malaria transmission relationship for the NHDSS (prepared for submission in *Trends in Parasitology*). In the last manuscript (chapter 5), predicted EIR for Manhiça HDSS were related to mortality data to estimate their effects (prepared for submission in *PLOS ONE*).

The methodology and detailed discussion of the findings are provided in each of the four chapters. This section reports a summary of the main findings, likely implications to malaria control interventions, limitation and conclusions.

The MTIMBA data were collected in 5 different countries with heterogeneous malaria exposures (Beier et al., 1999; Hay et al., 2000; Kelly-Hope and McKenzie, 2009). This is one of the largest entomological databases currently available that can be used to measure directly malaria transmission intensity and the continued effect of interventions on the disease. Data from all participating sites were collected under a similar protocol, and thus both within and between sites comparisons are possible.

In the 1990s, the South African Medical Research Council coordinated the Mapping Malaria Risk in Africa (MARA/ARMA, 1998) project which set up the first comprehensive malaria risk database across Africa (Le Sueur et al., 1997). The project however was based on published and unpublished data from parasite prevalence surveys. The MTIMBA project gathered large entomological data from West, East and Southern Africa (Figure 1.3) with high spatial and temporal resolutions, the characteristics that are relevant for assessing heterogeneity in disease transmission.

The project generated large geostatistical sparse data that were collected repeatedly over many locations as described in chapter 2. The longitudinal entomological data were influenced by seasonal climatic changes. Many locations were found with many zeros especially in the dry periods. The analysis developed zero-inflated Bayesian geostatistical models that improve the estimates of malaria exposure surfaces. Zero-inflated models have been applied mostly to count data (Amek et al., 2012; Soares Magalhães et al., 2011; Vounatsou et al., 2009) and least to binomial data (Amek et al., 2011). Motivated by the MTIMBA data, we developed and further extended analyses of the MTIMBA work using NHDSS and Manhiça data. The methodology for estimating malaria exposure surfaces using Navrongo and Manhica data is presented in chapters 3 and 5 respectively.

Previous studies have used the mathematical model of the Garki project (Gemperli et al., 2006) to convert age-specific prevalence at each location to a common scale of transmission intensity measure. The EIR was assumed to be constant over time period or season. The EIR estimates were log transformed to approximately achieve normality (Gemperli et al., 2006a; Gemperli et al., 2006b). The assumption of normality is rarely fulfilled especially because of the large number of zeros arising during the dry season. EIR arises from the product of binomial (sporozoite rate) and negative binomial (mosquito density) distributions. As indicated in chapters 3 and 5, EIR is estimated from models fitting separately the binomial sporozoite rate data and negative binomial count data. Bayesian kriging was then used to predict malaria transmission intensity at unsampled locations.

Improvements in geographic technologies such as Geographic Information Systems (GIS), Global Positioning Systems (GPS) and remote sensing have enabled an assimilation of climatic and environmental proxies at high space and time resolutions that can be used to predict health

outcomes. Malaria transmission intensity is driven by environmental and climatic factors (Kelly-Hope et al., 2009). In this analysis, the environmental and climatic covariates were extracted from satellite images at very high spatial and temporal resolutions. Bayesian geostatistical models were used to predict monthly malaria exposures surfaces at 250m by 250m spatial resolution for the two sites.

Mapping in malaria epidemiology is a useful tool especially in identifying potential foci of transmission to set priorities in terms of resource allocation and for assessing progress towards control program.

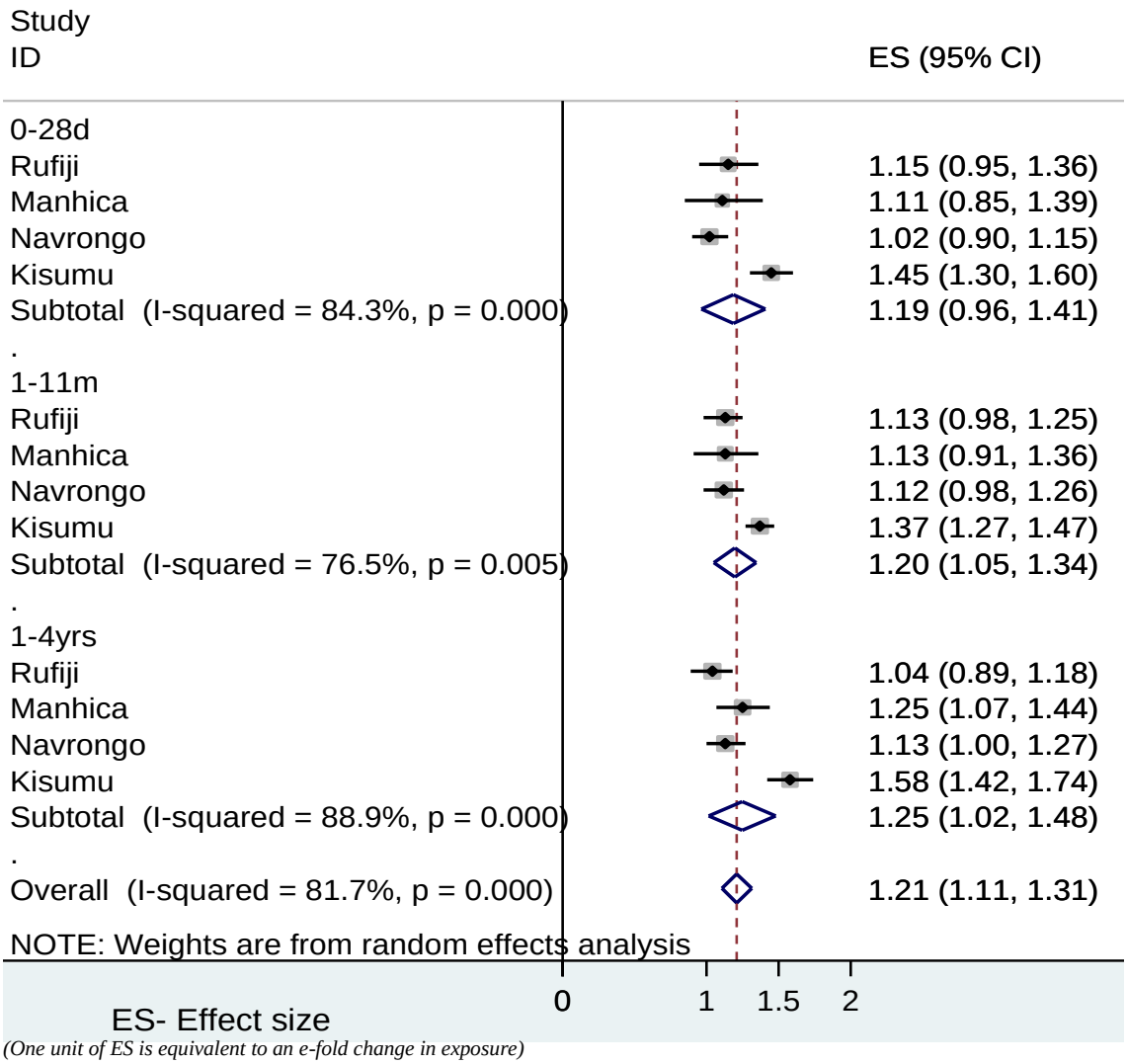
To our knowledge, there are no malaria risk maps specifically for the Navrongo HDSS. Even for the Manhica HDSS, the existing malaria risk maps were based on parasitological data that were generated by children under 10 years between 1996 and 1999 (Abellana et al., 2008; Escaramís et al., 2011).

Previous maps that were generated from the MARA\ARMA databases were based on a few data points (Craig et al., 1999; Gemperli et al., 2006; Gosoniu et al., 2006; Kazembe et al., 2006; Kleinschmidt et al., 2001). The databases consist of historical malaria prevalence surveys that are sparse and are unlikely to represent the current disease conditions. Other global malaria risk mapping attempts were also based on parasitological data (Gething et al., 2011; Hay and Snow, 2006). These data were collected at different seasons among different age groups making it difficult to adjust for such heterogeneity during the mapping process. A continental malaria endemicity levels map that was based on previous entomological surveys used data collected from only 15 countries (Hay et al., 2000). Chapters 3 and 5 used large amount of entomological data to derive more accurate malaria risk maps for the two HDSS in West and Southern Africa respectively. Climate and environmental predictors were also extracted at high spatial and

temporal resolutions. These maps can therefore be used to measure progress towards malaria elimination. Although monthly maps show changes in transmission intensity based on the climatic season, the geographical distribution remains constant (Figures 3.4 and 5.2). The findings therefore clearly showed that temporal heterogeneity is more important than the spatial variation in the two sites. The same situation was also observed from the analyses of the Rufiji and Kisumu MTIMBA-HDSS data (Amek et al., 2012; Rumisha, 2013) implying that the within site transmission is influenced mostly by ecological factors. Due to scarcity of resources, malaria control interventions such as insect treated-nets, indoor residual spraying and other vector control measures should target high transmission pockets especially towards the beginning of the rain season. . Although the spatial distribution of different diseases in the same area would have been helpful in identifying high risk clusters for integrated intervention approaches in order to improve health services delivery, the HDSS cover a smaller area. Implement malaria control interventions in an HDSS is not practically relevant. However, there is a need to understand the geographical distributions of malaria transmission intensity in these two sites in relation to health service delivery. High transmission areas are important for monitoring malaria early warning signs since the disease is mostly affected by seasonal changes.

The mortality-malaria transmission intensity relationship across all age groups was assessed in chapters 4 and 5 of the thesis. A positive log-linear relationship between mortality and EIR was observed in all age categories. Previous studies based on either meta-analysis (Ross and Smith, 2006; Smith et al., 2001; Snow and Marsh, 1995) or linking DHS mortality and the MARA databases (Gemperli et al, 2004) focused on children under five years and provided inconsistent results about this relationship. The observed positive effect of EIR on all-cause mortality among children < 5 years is consistent with the findings from Rufuji and Kisumu MTIMBA-HDSS

data (Rumisha 2013; Amek 2013). Preliminary meta-analysis using estimates generated through rigorous analyses of data from 4 MTIMBA sites reported an overall positive effect of EIR on mortality which is even statistically important (Figure 6.1). In these studies, the primary outcome was the frequency of death and the effects were therefore expressed as risk ratios.



(One unit of ES is equivalent to an e-fold change in exposure)

Figure 6. 1: Meta analysis for effect of EIR on all-cause mortality among children

Estimates from Bayesian meta-analysis formulation indicate that there is 99.9% posterior probability that mortality is associated with increased malaria transmission intensity among

children < 5. This implies measures aimed at reducing mosquito contacts in endemic areas will improve child survival in such areas. This suggests that acquired immunity which develops with repeated infections in these countries still occurs at an early age. Even with increased control interverntions in SSA, severe malaria cases and high malaria cause mortality are still common among young children (Abdullah et al., 2007; Carneiro et al., 2010; Roca-Feltrer et al., 2010). The established positive relationship between all-cause mortality and malaria exposure among children < 5 years might reflect the true situation in SSA. Further research however is needed to monitor this relationship in school age children and adult population. Meta analysis including all other age groups will be condicted when site-specific analyses are completed.

Use of ITN is one of the interventions that has been assocated with reduced mortality malaria endemic areas among children (Lengeler, 2004; Phillips-Howard et al., 2003; Steketee and Campbell, 2010). As indicate in chapter 4, ownership of nets was used as a proxy for use. After adjusting for spatio-temporal correlation, the data shows that children from compounds with nets received a protective effect. With the current scale-up of malaria control interventions, HDSS need to collect this information where possible at personal level.

6.2 Study limitations and challenges

The INDEPTH network malaria working group commissioned the MTIMBA project in early 2000 and the different sites completed data collection towards the end of 2004. Although site specific data were merged in 2005, analysis did not start till 2008. This was partly due to sites' lack of expertise especially in the field of statistics to handle the analysis. The first attempts to analyze MTIMBA data were intially done based on a sub-population of children under 15 years (Abdullah et al., 2007) and did not include statistical methods that take into consideration all the data characteristics. The analysis of MTIMBA data was therefore used as a platform to build

statastical, epidemiologal and writing capacity skills among African young scientists. This has been done through course works, seminars and mentorship. Currently, the analyses of the two MTIMBA sites data of Rufiji in Tanzania (Rumisha, 2013) and Kisumu in Kenya (Amek, 2013) were completed, and work for the remaining sites is still on-going. Despite all the delays, this rigorous training process will help HDSS not only to improve data collection and management skills but also carry out proper analysis for informed decision making.

Cause of death data were obtained using a modified verbal Autopsy (VA) tool derived from both, the World Health Organization (WHO) and the standard, site-specific VA questionnaires. At the end of the MTIMBA data collection period (end of 2004), complete VA data for NHDSS were not available. Manhica HDSS provided VA results of 1732 (85%) out of the 2027 total observed deaths. Although malaria mortality in SSA is common among the under fives (Becher et al., 2008c; Liu et al., 2012) , no deaths in this age category was assigned to the disease. The VA method is known to overestimate malaria deaths partly due to diseases with similar symptoms such as pneumonia and acute respiratory infections (Hammer et al., 2006). There is a need therefore to test the VA tool in Manhica HDSS against hospital data or models (Oti and Kyobutungi, 2010) in order to ascertain it's sensitivity and specificity.

Geostatistical analyis requires coordinates of all the units of analysis. Like other MTIMBA sites, the NHDSS and Manhica HDSS lacked data coordinates for all the compounds/households. Coordinates were also necessary to allign all the other factors. Lack of coordinates forced us to drop large amount of data generated by compounds that were not georeferenced. Similarly, while estimating the spatial process for mortality-malaria transmission models, the spatial random effect was assigned at village level. However, village maps were not available to ascertain the distribution of compounds and thier central positions. Instead, an average of the

georeferenced compounds within each village was used to generated the centroids. Accurate and complete georeferenced data is important for modeling and for determining the disease risk. The HDSS need to apply the current improvements in geographic technologies to map all their compounds and villages.

In the process of developing malaria transmission models, climate and environmental predictors used were extracted from remote sensing data at different spatio-temporal resolutions. Each of the covariates at particular location was summarized into a single value (either mean or total) over a specified period of time defined by either one, two or three months before the actual survey time. The mosquioto data on the other hand that were obtained every fortnight at sampled locations were linked to the summaries of environmental and climatic data before analysis. Using aggregated temporal covariates is likely to influence the accuracy of model parameters. This assumes that climate and environment predictors of an area are constant throughout the specified time interval which is not true.

While modeling malaria transmission intensity, lag time analyses were performed to ascertain a suitable combiation of predictors to both mosquito density and sporozoite rata data. Covariates were selected using standard negative binomial and logistic regression models. The best combiation of predictors was selected basing on Akaike information creterion (AIC) in additon to a comparision between fitted and predicted values. Covariates selected using standard regression models ignores spatial correlation that exists in the entomological data. Bayesian variable selection is is a relevant to the MTIMBA. This method identify the best fit of covariates from a model with highest predictive ability (Dellaportas et al., 2002). Although Bayesian variable selection method has previously been applied to malaria risk mapping (Giardina et al., 2012) it is computationally expensive when applied to large data.

Mortality models used EIR covariate and individual age. Age was measured either in days, months or years. Studying malaria epidemiology is broad and requires all the characteristics that influence it. These include transmission intensity which was measured by EIR, treatment regimen, people's behaviours and malaria control interventions in place. The effect of such factors is important in designing and evaluating malaria control interventions bacause the disease is always driven by local conditions.

6.3 Conclusion

This analysis reported a positive association between all-cause mortality and malaria transmission intensity in all age categories. Such relationships especially among the adults should be interpreted carefully. People in endemic areas develop clinical immunity at an early age and severe malaria cases usually reduce with increasing age. Further site-specific analyses followed by meta-analysis will useful in providing reliable evidence about mortality-malaria transmission relationship across different age groups in SSA.

Heterogeneity in malaria transmission intensity mostly driven by temporal changes was observed in the two MTIMBA sites. Spatial variation was not so strong because of the small sizes of these HDSS. There is a need therefore for HDSS to target their interventions especially in the transmission peak seasons in order to avoid disease epidemics.

Bayesian geostatistical models developed under the MTIMBA project can be adopted and used for the analyses of other HDSS data with similar characteristics namely; large data, longitudinal in nature, seasonal and zero-inflated. However such models are applicable to data collected at the same spatial and temporal ressolution.

Bibliography

Abdullah, S., Adazu, K., Masanja, H., Diallo, D., Hodgson, A., Ilboudo-Sanogo, E., Nhacolo, A., Owusu-Agyei, S., Thompson, R., Smith, T., Binka, F.N., 2007. Patterns of Age-Specific Mortality in Children in Endemic Areas of Sub-Saharan Africa. Am J Trop Med Hyg 77, 99–105.

Abellana, R., Ascaso, C., Aponte, J., Saute, F., Nhalungo, D., Nhacolo, A., Alonso, P., 2008. Spatio-seasonal modeling of the incidence rate of malaria in Mozambique. Malaria Journal 7, 228.

Abonuusum, A., Owusu-Daako, K., Tannich, E., May, J., Garms, R., Kruppa, T., 2010. Malaria transmission in two rural communities in the forest zone of Ghana. Parasitol Res.

Adazu, K., Lindblade, K.A., Rosen, D.H., Odhiambo, F., Ofware, P., Kwach, J., Van Eijk, A.M., Decock, K.M., Amornkul, P., Karanja, D., Vulule, J.M., Slutsker, L., 2005. Health and demographic surveillance in rural western Kenya: a platform for evaluating interventions to reduce morbidity and mortality from infectious diseases. Am. J. Trop. Med. Hyg 73, 1151–1158.

Adjuik, M., Kanyomse, E., Kondayire, F., Wak, G., Hodgson, A., 2010. Clustering of under-five mortality in the Navrongo HDSS in the Kassena-Nankana District of northern Ghana. Global Health Action 3.

Adjuik, M., Smith, T., Clark, S., Todd, J., Garrib, A., Kinfu, Y., Kahn, K., Mola, M., Ashraf, A., Masanja, H., Adazu, K., Adazu, U., Sacarlal, J., Alam, N., Marra, A., Gbangou, A., Mwageni, E., Binka, F., 2006. Cause-specific mortality rates in sub-Saharan Africa and Bangladesh. Bull. World Health Organ. 84, 181–188.

Akachi, Y., Atun, R., 2011. Effect of Investment in Malaria Control on Child Mortality in Sub-Saharan Africa in 2002–2008. PLoS ONE 6, e21309.

Allison, D.P., 1982. Discrete-time methods for the analysis of histories. Sociological Methodology 13.

Alonso, P.L., Lindsay, S.W., Armstrong, J.R., Conteh, M., Hill, A.G., David, P.H., Fegan, G., de Francisco, A., Hall, A.J., Shenton, F.C., 1991. The effect of insecticide-treated bed nets on mortality of Gambian children. Lancet 337, 1499–1502.

Amek, N., Bayoh, N., Hamel, M., Lindblade, K.A., Gimnig, J., Laserson, K.F., Slutsker, L., Smith, T., 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, 283–290.

Amek, N., Bayoh, N., Hamel, M., Lindblade, K.A., Gimnig, J., Odhiambo, F., Laserson, K.F., Slutsker, L., Smith, T., Vounatsou, P., 2012. Spatial and temporal dynamics of malaria transmission in rural Western Kenya. Parasites & vectors 5, 86.

Amek, N.., 2013. Bayesian spatio-temporal modelling of the relationship between mortility and malaria transmission in rural western kenya. (Thesis submitted for PhD, University of Basel).

Appawu, M., Owusu-Agyei, S., Dadzie, S., Asoala, V., Anto, F., Koram, K., Rogers, W., Nkrumah, F., Hoffman, S.L., Fryauff, D.J., 2004. Malaria transmission dynamics at a site in northern Ghana proposed for testing malaria vaccines. Trop. Med. Int. Health 9, 164–170.

Aranda, C, Aponte, J.J., Saute, F., Casimiro, S., Pinto, J., Sousa, C., Rosario, V.D., Petrarca, V., Dgedge, M., Alonso, P., 2005. Entomological characteristics of malaria transmission in Manhiça, a rural area in southern Mozambique. J. Med. Entomol. 42, 180–186.

Arrow, K.J., Panosian, C., Gelband, H., 2004. Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance [WWW Document]. URL http://www.nap.edu/openbook.php?record_id=11017&page=136 (accessed 5.11.12).

Ashton, R.A., Kefyalew, T., Tesfaye, G., Pullan, R.L., Yadeta, D., Reithinger, R., Kolaczinski, J.H., Brooker, S., 2011. School-based surveys of malaria in Oromia Regional State, Ethiopia: a rapid survey method for malaria in low transmission settings. Malaria Journal 10, 25.

Baker, W.E., Dempster, T.E., Yule, H., 1868. The Prevalence of organic disease of the spleen as a test for detecting malarious localities in hot climates: being the report of a committee assembled by general orders Commander-in-Chief, dated the 16th September 1854. Office of the Superintendent of Government Printing.

Banerjee, S., Gelfand, A.E., Carlin, B.P., 2003. Hierarchical Modeling and Analysis for Spatial Data, 1st ed. Chapman and Hall/CRC.

Banerjee, S., Gelfand, A.E., Finley, A.O., Sang, H., 2008. Gaussian predictive process models for large spatial data sets. J R Stat Soc Series B Stat Methodol 70, 825–848.

Bardají, A., Sigauque, B., Sanz, S., Maixenchs, M., Ordi, J., Aponte, J.J., Mabunda, S., Alonso, P.L., Menéndez, C., 2011. Impact of malaria at the end of pregnancy on infant mortality and morbidity. J. Infect. Dis. 203, 691–699.

Baume, C.A., Marin, M.C., 2007. Intra-household mosquito net use in Ethiopia, Ghana, Mali, Nigeria, Senegal, and Zambia: are nets being used? Who in the household uses them? Am. J. Trop. Med. Hyg. 77, 963–971.

Becher, H., Kynast-Wolf, G., Sié, A., Ndugwa, R., Ramroth, H., Kouyaté, B., Müller, O., 2008. Patterns of malaria: cause-specific and all-cause mortality in a malaria-endemic area of west Africa. Am. J. Trop. Med. Hyg. 78, 106–113.

Beier, J.C., 1998. Malaria Parasite Development in Mosquitoes. Annual Review of Entomology 43, 519–543.

Beier, J.C., Killeen, G.F., Githure, J.I., 1999. Short report: entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa. Am. J. Trop. Med. Hyg. 61, 109–113.

Binka, F.N., Bawah, A.A., Phillips, J.F., Hodgson, A., Adjuik, M., MacLeod, B., 2007. Rapid achievement of the child survival millennium development goal: evidence from the Navrongo experiment in Northern Ghana. Tropical Medicine & International Health 12, 578–593.

Binka, F.N., Indome, F., Smith, T., 1998. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. The American Journal of Tropical Medicine and Hygiene 59, 80 –85.

Binka, F.N., Kubaje, A., Adjuik, M., Williams, L.A., Lengeler, C., Maude, G.H., Armah, G.E., Kajihara, B., Adiamah, J.H., Smith, P.G., 1996. Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. Trop. Med. Int. Health 1, 147–154.

Binka, F.N., Morris, S.S., Ross, D.A., Arthur, P., Aryeetey, M.E., 1994. Patterns of malaria morbidity and mortality in children in northern Ghana. Trans. R. Soc. Trop. Med. Hyg. 88, 381–385.

Binka, F.N., Ngom, P., Phillips, J.F., Adazu, K., MacLEOD, B.B., 1999. Assessing population dynamics in a rural African society: The Navrongo demographic surveillance system Journal of Biosocial Science 31, 375–391.

Black, R.E., Cousens, S., Johnson, H.L., Lawn, J.E., Rudan, I., Bassani, D.G., Jha, P., Campbell, H., Walker, C.F., Cibulskis, R., Eisele, T., Liu, L., Mathers, C., 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet 375, 1969–1987.

Blacker, J., 2004. The impact of AIDS on adult mortality: evidence from national and regional statistics. AIDS 18 Suppl 2, S19–26.

Brezger, A., Kneib, T., Lang, S., 2005. BayesX: Analyzing Bayesian Structural Additive Regression Models. Journal Of Statistical Software 14, 1–22.

Briët, O.J.T., Vounatsou, P., Gunawardena, D.M., Galappaththy, G.N.L., Amerasinghe, P.H., 2008. Temporal correlation between malaria and rainfall in Sri Lanka. Malar. J. 7, 77.

Bryce, J., Boschi-Pinto, C., Shibuya, K., Black, R.E., 2005. WHO estimates of the causes of death in children. The Lancet 365, 1147–1152.

Byass, P., Berhane, Y., Emmelin, A., Kebede, D., Andersson, T., Högberg, U., Wall, S., 2002. The role of demographic surveillance systems (DSS) in assessing the health of communities: An example from rural Ethiopia. Public Health 116, 145–150.

Carneiro, I., Roca-Feltrer, A., Griffin, J.T., Smith, L., Tanner, M., Schellenberg, J.A., Greenwood, B., 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, e8988.

Carter, R., Mendis, K.N., Roberts, D., 2000. Spatial targeting of interventions against malaria. Bull World Health Organ 78, 1401–1411.

Chapter 22. Navrongo DSS Ghana [WWW Document], 2002. URL http://www.idrc.ca/en/ev-43032-201-1-DO_TOPIC.html (accessed 3.15.11).

Charlwood, J.D., Kihonda, J., Sama, S., Billingsley, P.F., Hadji, H., Verhave, J.P., Lyimo, E., Luttikhuizen, P.C., Smith, T., 1995. The Rise and Fall of Anopheles Arabiensis (Diptera: Culicidae) in a Tanzanian Village. Bulletin of Entomological Research 85, 37–44.

Chase, C., Sicuri, E., Sacoor, C., Nhalungo, D., Nhacolo, A., Alonso, P.L., Menéndez, C., 2009. Determinants of household demand for bed nets in a rural area of southern Mozambique. Malaria Journal 8, 132.

Christopher, J.B., Le May, A., Lewin, S., Ross, D.A., 2011. Thirty years after Alma-Ata: a systematic review of the impact of community health workers delivering curative interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in sub-Saharan Africa. Hum Resour Health 9, 27.

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, e1001142.

Clements, A.C.A., Moyeed, R., Brooker, S., 2006. Bayesian geostatistical prediction of the intensity of infection with Schistosoma mansoni in East Africa. Parasitology 133, 711–719.

Clerk, C.A., Bruce, J., Greenwood, B., Chandramohan, D., 2009. The epidemiology of malaria among pregnant women attending antenatal clinics in an area with intense and highly seasonal malaria transmission in northern Ghana. Trop. Med. Int. Health 14, 688–695.

Coetzee, M., Craig, M., le Sueur, D., 2000. Distribution of African malaria mosquitoes belonging to the Anopheles gambiae complex. Parasitol. Today (Regul. Ed.) 16, 74–77.

Coetzee, M., Fontenille, D., 2004. Advances in the study of Anopheles funestus, a major vector of malaria in Africa. Insect Biochemistry and Molecular Biology 34, 599–605.

Collins, W.E., 2012. Plasmodium knowlesi: A Malaria Parasite of Monkeys and Humans*. Annual Review of Entomology 57, 107–121.

Cox, D., 1972. Regression Models and Life-Tables. Journal of the Royal Statistical Society. Series B (Methodological) 34, 187–220.

Craig, M.H., Snow, R.W., le Sueur, D., 1999. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitol. Today (Regul. Ed.) 15, 105–111.

Cressie, N., 1993. Statistics for Spatial Data, Revised Edition. ed. Wiley-Interscience.

D'Alessandro, U., Olaleye, B.O., McGuire, W., Langerock, P., Bennett, S., Aikins, M.K., Thomson, M.C., Cham, M.K., Cham, B.A., Greenwood, B.M., 1995. Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. Lancet 345, 479–483.

De Savigny, D., Mayombana, C., Mwageni, E., Masanja, H., Minhaj, A., Mkilindi, Y., Mbuya, C., Kasale, H., Reid, G., 2004. Care-seeking patterns for fatal malaria in Tanzania. Malaria Journal 3, 27.

De Souza, D., Kelly-Hope, L., Lawson, B., Wilson, M., Boakye, D., 2010. Environmental factors associated with the distribution of Anopheles gambiae s.s in Ghana; an important vector of lymphatic filariasis and malaria. PLoS ONE 5, e9927.

Dellaportas, P., Forster, J., Ntzoufras, I., 2002. On Bayesian model and variable selection using MCMC. Statistics and Computing 12, 27–36.

Deressa, W., Fantahun, M., Ali, A., 2007. Malaria-related mortality based on verbal autopsy in an area of low endemicity in a predominantly rural population in Ethiopia. Malar. J 6, 128.

Dery, D.B., Brown, C., Asante, K.P., Adams, M., Dosoo, D., Amenga-Etego, S., Wilson, M., Chandramohan, D., Greenwood, B., Owusu-Agyei, S., 2010. Patterns and seasonality of malaria transmission in the forest-savannah transitional zones of Ghana. Malar. J. 9, 314.

Deville, J.-C., Tillé, Y., 2004. Efficient Balanced Sampling: The Cube Method. Biometrika 91, 893–912.

Diallo, D.A., Cousens, S.N., Cuzin-Ouattara, N., Nebié, I., Ilboudo-Sanogo, E., Esposito, F., 2004. Child mortality in a West African population protected with insecticide-treated curtains for a period of up to 6 years. Bull World Health Organ 82, 85–91.

Diggle, P.J., Moyeed, R.A., Tawn, J.A., 1998. Model-based Geostatistics. APPLIED STATISTICS 47, 299–350.

Drakeley, C., Schellenberg, D., Kihonda, J., Sousa, C.A., Arez, A.P., Lopes, D., Lines, J., Mshinda, H., Lengeler, C., Schellenberg, J.A., Tanner, M., Alonso, P., 2003. An estimation of the entomological inoculation rate for Ifakara: a semi-urban area in a region of intense malaria transmission in Tanzania. Tropical Medicine & International Health 8, 767–774.

Drakeley, Chris J, Carneiro, I., Reyburn, H., Malima, R., Lusingu, J.P.A., Cox, J., Theander, T.G., Nkya, W.M.M.M., Lemnge, M.M., Riley, E.M., 2005a. Altitude-dependent and -independent variations in Plasmodium falciparum prevalence in northeastern Tanzania. J. Infect. Dis. 191, 1589–1598.

Drakeley, C J, Corran, P.H., Coleman, P.G., Tongren, J.E., McDonald, S.L.R., Carneiro, I., Malima, R., Lusingu, J., Manjurano, A., Nkya, W.M.M., Lemnge, M.M., Cox, J., Reyburn, H., Riley, E.M., 2005b. Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure. Proc. Natl. Acad. Sci. U.S.A. 102, 5108–5113.

Ecker, M.D., Gelfand, A.E., 1997. Bayesian Variogram Modeling for an Isotropic Spatial Process. Journal of Agricultural, Biological, and Environmental Statistics 2, 347–369.

Eisele, T.P., Larsen, D., Steketee, R.W., 2010. Protective efficacy of interventions for preventing malaria mortality in children in Plasmodium falciparum endemic areas. International Journal of Epidemiology 39, i88–i101.

Escaramís, G., Carrasco, J.L., Aponte, J.J., Nhalungo, D., Nhacolo, A., Alonso, P., Ascaso, C., 2011. Spatio-temporal analysis of mortality among children under the age of five in Manhiça (Mozambique) during the period 1997-2005. Int J Health Geogr 10, 14.

Etard, J.-F., Le Hesran, J.-Y., Diallo, A., Diallo, J.-P., Ndiaye, J.-L., Delaunay, V., 2004. Childhood mortality and probable causes of death using verbal autopsy in Niakhar, Senegal, 1989–2000. International Journal of Epidemiology 33, 1286 –1292.

Finley, A., Banerjee, S., Carlin, B., 2007. spBayes: An R Package for Univariate and Multivariate Hierarchical Point-referenced Spatial Models.

Finley, A.O., Sang, H., Banerjee, S., Gelfand, A.E., 2009. Improving the performance of predictive process modeling for large datasets. Comput Stat Data Anal 53, 2873–2884.

Gelfand, A., Smith, A., 1990. Sampling-Based Approaches to Calculating Marginal Densities. Journal of the American Statistical Association 85, 398–409.

Gelfand, A.E., 2007. Hierarchical Modeling and Analysis for Spatial Data, 1st ed. CRC Press.

Gemperli, A., 2003. Development of Spatial statisticam methods for modeling point referenced spatial data in malaria epidemiology. (PhD Thesis).

Gemperli, Armin, Sogoba, N., Fondjo, E., Mabaso, M., Bagayoko, M., Briët, O.J.T., Anderegg, D., Liebe, J., Smith, T., Vounatsou, P., 2006a. Mapping malaria transmission in West and Central Africa. Trop. Med. Int. Health 11, 1032–1046.

Gemperli, A., Vounatsou, P., Kleinschmidt, I., Bagayoko, M., Lengeler, C., Smith, T., 2004. Spatial Patterns of Infant Mortality in Mali: The Effect of Malaria Endemicity. American Journal of Epidemiology 159, 64 –72.

Gemperli, A, Vounatsou, P., Sogoba, N., Smith, T., 2006b. Malaria mapping using transmission models: application to survey data from Mali. Am. J. Epidemiol. 163, 289–297.

Gething, P.W., Patil, A.P., Smith, D.L., Guerra, C.A., Elyazar, I.R.F., Johnston, G.L., Tatem, A.J., Hay, S.I., 2011. A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar. J. 10, 378.

Ghana VAST Study Team, 1993. Vitamin A supplementation in northern Ghana: effects on clinic attendances, hospital admissions, and child mortality. Lancet 342, 7–12.

GHS NMCP-Report, 2008. National malaria control programme annual report 2008.

Giardina, F., Gosoniu, L., Konate, L., Diouf, M.B., Perry, R., Gaye, O., Faye, 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, e32625.

Gillies, M., Mielion, B., 1968. The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). Publications of the South African Institute for Medical Research 54, 1–343.

Githinji, S., Herbst, S., Kistemann, T., Noor, A.M., 2010. Mosquito nets in a rural area of Western Kenya: ownership, use and quality. Malar. J. 9, 250.

Gosoniu D., Vounatsou P., Kahn K.,, Tille Y., 2011. Geostatistical modeling of large non-Gaussian irregularly distributed data. Journal of Computational Statistics and Data Analysis (in press).

Gosoniu, L., Veta, A.M., Vounatsou, P., 2010. Bayesian geostatistical modeling of Malaria Indicator Survey data in Angola. PLoS ONE 5, e9322.

Gosoniu, L., Vounatsou, P., 2011. Non-stationary partition modeling of geostatistical data for malaria risk mapping. Journal of Applied Statistics 38, 3–13.

Gosoniu, L., Vounatsou, P., Sogoba, N., Smith, T., 2006. Bayesian modelling of geostatistical malaria risk data. Geospat Health 1, 127–139.

Greenwood, B., Mutabingwa, T., 2002. Malaria in 2002. Nature 415, 670–672.

GSS, GHS, 2009. Ghana demographic and health survey 2008.

Habluetzel, A., Diallo, D.A., Esposito, F., Lamizana, L., Pagnoni, F., Lengeler, C., Traoré, C., Cousens, S.N., 1997. Do insecticide-treated curtains reduce all-cause child mortality in Burkina Faso? Trop. Med. Int. Health 2, 855–862.

Hall, D.B., 2000. Zero-inflated Poisson and binomial regression with random effects: a case study. Biometrics 56, 1030–1039.

Hamad, A.A., Nugud, A.E.H.D., Arnot, D.E., Giha, H.A., Abdel-Muhsin, A.-M.A., Satti, G.M.., Theander, T.G., Creasey, A.M., Babiker, H.A., Elnaiem, D.-E.A., 2002. A marked seasonality of malaria transmsission in two rural sites in eastern Sudan. Acta Tropica 83, 71–82.

Hammer, G.P., Somé, F., Müller, O., Kynast-Wolf, G., Kouyaté, B., Becher, H., 2006. Pattern of cause-specific childhood mortality in a malaria endemic area of Burkina Faso. Malar. J. 5, 47.

Harbach, R.E., 2004. The classification of genus Anopheles (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. Bull. Entomol. Res. 94, 537–553.

Hay, J L, Pettitt, A.N., 2001. Bayesian analysis of a time series of counts with covariates: an application to the control of an infectious disease. Biostatistics 2, 433–444.

Hay, S.I., Guerra, C.A., Gething, P.W., Patil, A.P., Tatem, A.J., Noor, A.M., Kabaria, C.W., Manh, B.H., Elyazar, I.R.F., Brooker, S., Smith, D.L., Moyeed, R.A., Snow, R.W., 2009. A World Malaria Map: Plasmodium falciparum Endemicity in 2007. PLoS Med 6, e1000048.

Hay, S.I., Okiro, E.A., Gething, P.W., Patil, A.P., Tatem, A.J., Guerra, C.A., Snow, R.W., 2010. Estimating the global clinical burden of Plasmodium falciparum malaria in 2007. PLoS Med. 7, e1000290.

Hay, S.I., Rogers, D.J., Toomer, J.F., Snow, R.W., 2000. Annual Plasmodium falciparum entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review. Trans. R. Soc. Trop. Med. Hyg. 94, 113–127.

Hay, S.I., Smith, D.L., Snow, R.W., 2008. Measuring malaria endemicity from intense to interrupted transmission. Lancet Infect Dis 8, 369–378.

Hay, S.I., Snow, R.W., 2006. The Malaria Atlas Project: Developing Global Maps of Malaria Risk. PLoS Med 3, e473.

Himeidan, Y.E., Elzaki, M.M., Kweka, E.J., Ibrahim, M., Elhassan, I.M., 2011. Pattern of malaria transmission along the Rahad River basin, Eastern Sudan. Parasit Vectors 4, 109.

Hoeting, J.A., Davis, R.A., Merton, A.A., Thompson, S.E., 2006. Model selection for geostatistical models. Ecol Appl 16, 87–98.

Houweling, T.A.J., Kunst, A.E., Moser, K., Mackenbach, J.P., 2006. Rising under-5 mortality in Africa: who bears the brunt? Trop. Med. Int. Health 11, 1218–1227.

Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin distance designs. Journal of Statistical Planning and Inference 26, 131–148.

Joshi, R., Lopez, A.D., MacMahon, S., Reddy, S., Dandona, R., Dandona, L., Neal, B., 2009. Verbal autopsy coding: are multiple coders better than one? Bull. World Health Organ. 87, 51–57.

Kasasa, S., et al. The effect of malaria transmission intensity on mortality in Manhica Demographic surveillance site, southern Mozambique. (in Proparation)

Kasasa, S., Asoala, V., Gosoniu, L., Anto, F., Adjuik, M., Tindana, C., Smith, T., Owusu-Agyei, S., Vounatsou, P., 2013. Spatio-temporal malaria transmission patterns in Navrongo demographic surveillance site, northern Ghana. Malar. J. 12, 63.

Kasasa, S., Rumisha, S., Amek, N.., Diboulo, E., Huho, B., Smith, T., VOUNATSOU, P.. Malaria transmission intensity amd mortality burden across Africa project: statistical issues and approaches to data analysis. (in Proparation).

Kazembe, Lawrence N, Kleinschmidt, I., Holtz, T.H., Sharp, B.L., 2006. Spatial analysis and mapping of malaria risk in Malawi using point-referenced prevalence of infection data. Int J Health Geogr 5, 41.

Kelly-Hope, L.A., Hemingway, J., McKenzie, F.E., 2009. Environmental factors associated with the malaria vectors Anopheles gambiae and Anopheles funestus in Kenya. Malar. J. 8, 268.

Kelly-Hope, L.A., McKenzie, F.E., 2009. The multiplicity of malaria transmission: a review of entomological inoculation rate measurements and methods across sub-Saharan Africa. Malar. J. 8, 19.

Kevin Baird, J., Bangs, M.J., Maguire, J.D., Barcus, M.J., 2002. Epidemiological Measures of Risk of Malaria, in: Doolan, D.L., Walker, J.M. (Eds.), Malaria Methods and Protocols, Methods in Molecular Medicine. Humana Press, pp. 13–22.

Kleinschmidt, I., Omumbo, J., Briët, O., van de Giesen, N., Sogoba, N., Mensah, N.K., Windmeijer, P., Moussa, M., Teuscher, T., 2001. An empirical malaria distribution map for West Africa. Trop. Med. Int. Health 6, 779–786.

Kloke, R.G., Nhamahanga, E., Hunt, R.H., Coetzee, M., 2011. Vectorial status and insecticide resistance of Anopheles funestus from a sugar estate in southern Mozambique. Parasit Vectors 4, 16.

Konaté, A.T., Yaro, J.B., Ouédraogo, A.Z., Diarra, A., Gansané, A., Soulama, I., Kangoyé, D.T., Kaboré, Y., Ouédraogo, E., Ouédraogo, A., Tiono, A.B., Ouédraogo, I.N., Chandramohan, D., Cousens, S., Milligan, P.J., Sirima, S.B., Greenwood, B., Diallo, D.A., 2011. Intermittent Preventive Treatment of Malaria Provides Substantial Protection against Malaria in Children Already Protected by an Insecticide-Treated Bednet in Burkina Faso: A Randomised, Double-Blind, Placebo-Controlled Trial. PLoS Med 8, e1000408.

Koram, K.A., Owusu-Agyei, S., Utz, G., Binka, F.N., Baird, J.K., Hoffman, S.L., Nkrumah, F.K., 2000. Severe anemia in young children after high and low malaria transmission seasons in the Kassena-Nankana district of northern Ghana. Am. J. Trop. Med. Hyg. 62, 670–674.

Korenromp, E.L., Miller, J., Cibulskis, R.E., Kabir Cham, M., Alnwick, D., Dye, C., 2003. Monitoring mosquito net coverage for malaria control in Africa: possession vs. use by children under 5 years. Trop. Med. Int. Health 8, 693–703.

Kristan, M., Abeku, T.A., Beard, J., Okia, M., Rapuoda, B., Sang, J., Cox, J., 2008. Variations in entomological indices in relation to weather patterns and malaria incidence in East African highlands: implications for epidemic prevention and control. Malar. J. 7, 231.

Lambert, D., 1992. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. Technometrics 34, 1–14.

Le Sueur, D., Binka, F., Lengeler, C., De Savigny, D., Snow, B., Teuscher, T., Toure, Y., 1997. An atlas of malaria in Africa. Afr Health 19, 23–24.

Lengeler, C., 2004. Insecticide-treated bed nets and curtains for preventing malaria. Cochrane Database Syst Rev CD000363.

Lengeler, C., Armstrong-Schellenberg, J., D'Alessandro, U., Binka, F., Cattani, J., 1998. Relative versus absolute risk of dying reduction after using insecticide-treated nets for malaria control in Africa. Trop. Med. Int. Health 3, 286–290.

Lengeler, C., Smith, T.A., Armstrong Schellenberg, J., 1997. Focus on the effect of bednets on malaria morbidity and mortality. Parasitol. Today (Regul. Ed.) 13, 123–124; author reply 125–126.

Lin, X., Wahba, G., Xiang, D., Gao, F., Klein, R., Klein, B., 2000. Smoothing Spline ANOVA Models for Large Data Sets with Bernoulli Observations and the Randomized GACV. The Annals of Statistics 28, 1570–1600.

Lindblade, K.A., Eisele, T.P., Gimnig, J.E., Alaii, J.A., Odhiambo, F., ter Kuile, F.O., Hawley, W.A., Wannemuehler, K.A., Phillips-Howard, P.A., Rosen, D.H., Nahlen, B.L., Terlouw, D.J., Adazu, K., Vulule, J.M., Slutsker, L., 2004. Sustainability of reductions in malaria transmission and infant mortality in western Kenya with use of insecticide-treated bednets: 4 to 6 years of follow-up. JAMA 291, 2571–2580.

Lindsay, S.W., Parson, L., Thomas, C.J., 1998. Mapping the ranges and relative abundance of the two principal African malaria vectors, Anopheles gambiae sensu stricto and An. arabiensis, using climate data. Proc. Biol. Sci. 265, 847–854.

Lines, J. d., Curtis, C. f., Wilkes, T. j., Njunwa, K. j., 1991. Monitoring Human-Biting Mosquitoes (Diptera: Culicidae) in Tanzania with Light-Traps Hung Beside Mosquito Nets. Bulletin of Entomological Research 81, 77–84.

Liu, L., Johnson, H.L., Cousens, S., Perin, J., Scott, S., Lawn, J.E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C., Black, R.E., 2012. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet 379, 2151–2161.

Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T., Murray, C.J.L., 2006. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 367, 1747–1757.

Lunn, D., Spiegelhalter, D., Thomas, A., Best, N., 2009. The BUGS project: Evolution, critique and future directions. Stat Med 28, 3049–3067.

Lusingu, J.P.A., Vestergaard, L.S., Mmbando, B.P., Drakeley, C.J., Jones, C., Akida, J., Savaeli, Z.X., Kitua, A.Y., Lemnge, M.M., Theander, T.G., 2004. Malaria morbidity and immunity among residents of villages with different Plasmodium falciparum transmission intensity in North-Eastern Tanzania. Malar. J. 3, 26.

Mabaso, M.L.H., Craig, M., Ross, A., Smith, T., 2007. Environmental predictors of the seasonality of malaria transmission in Africa: the challenge. Am. J. Trop. Med. Hyg. 76, 33–38.

Mabaso, M.L.H., Sharp, B., Lengeler, C., 2004. Historical review of malarial control in southern African with emphasis on the use of indoor residual house-spraying. Trop. Med. Int. Health 9, 846–856.

Manda, S., Meyer, R., 2005. Age at first marriage in Malawi: a Bayesian multilevel analysis using a discrete time-to-event model. Journal of the Royal Statistical Society: Series A (Statistics in Society) 168, 439–455.

Mathers, C.D., Fat, D.M., Inoue, M., Rao, C., Lopez, A.D., 2005. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull. World Health Organ. 83, 171–177.

Maxwell, C.A., Rwegoshora, R.T., Magesa, S.M., Curtis, C.F., 2006. Comparison of coverage with insecticide-treated nets in a Tanzanian town and villages where nets and insecticide are either marketed or provided free of charge. Malar. J. 5, 44.

Mbogo, C.M., Mwangangi, J.M., Nzovu, J., Gu, W., Yan, G., Gunter, J.T., Swalm, C., Keating, J., Regens, J.L., Shililu, J.I., Githure, J.I., Beier, J.C., 2003. Spatial and temporal heterogeneity of Anopheles mosquitoes and Plasmodium falciparum transmission along the Kenyan coast. Am. J. Trop. Med. Hyg 68, 734–742.

Mendis, C., Jacobsen, J.L., Gamage-Mendis, A., Bule, E., Dgedge, M., Thompson, R., Cuamba, N., Barreto, J., Begtrup, K., Sinden, R.E., Høgh, B., 2000. Anopheles arabiensis and An. funestus are equally important vectors of malaria in Matola coastal suburb of Maputo, southern Mozambique. Med. Vet. Entomol. 14, 171–180.

Menéndez, C., Bardají, A., Sigauque, B., Sanz, S., Aponte, J.J., Mabunda, S., Alonso, P.L., 2010. Malaria prevention with IPTp during pregnancy reduces neonatal mortality. PLoS ONE 5, e9438.

Minakawa, N., Sonye, G., Mogi, M., Githeko, A., Yan, G., 2002. The Effects of Climatic Factors on the Distribution and Abundance of Malaria Vectors in Kenya. Journal of Medical Entomology 39, 833–841.

Morris, S.S., Black, R.E., Tomaskovic, L., 2003. Predicting the distribution of under-five deaths by cause in countries without adequate vital registration systems. International Journal of Epidemiology 32, 1041 –1051.

Mukanga, D., Tibenderana, J.K., Peterson, S., Pariyo, G.W., Kiguli, J., Waiswa, P., Babirye, R., Ojiambo, G., Kasasa, S., Pagnoni, F., Kallander, K., 2012. Access, acceptability and utilization of community health workers using diagnostics for case management of fever in Ugandan children: a crosssectional study. Malaria journal 11, 121.

Munguambe, K., Pool, R., Montgomery, C., Bavo, C., Nhacolo, A., Fiosse, L., Sacoor, C., Nhalungo, D., Mabunda, S., Macete, E., Alonso, P., 2011. What drives community adherence to indoor residual spraying (IRS) against malaria in Manhiça district, rural Mozambique: a qualitative study. Malaria Journal 10, 344.

Murray, C.J.L., Rosenfeld, L.C., Lim, S.S., Andrews, K.G., Foreman, K.J., Haring, D., Fullman, N., Naghavi, M., Lozano, R., Lopez, A.D., 2012. Global malaria mortality between 1980 and 2010: a systematic analysis. Lancet 379, 413–431.

Murray, C.K., Gasser, R.A., Jr, Magill, A.J., Miller, R.S., 2008. Update on rapid diagnostic testing for malaria. Clin. Microbiol. Rev. 21, 97–110.

Muturi, E.J., Burgess, P., Novak, R.J., 2008. Malaria vector management: where have we come from and where are we headed? Am. J. Trop. Med. Hyg. 78, 536–537.

Nakamura, H., Ikeda, N., Stickley, A., Mori, R., Shibuya, K., 2011. Achieving MDG 4 in sub-Saharan Africa: what has contributed to the accelerated child mortality decline in Ghana? PLoS ONE 6, e17774.

Nevill, C.G., Some, E.S., Mung'ala, V.O., Mutemi, W., New, L., Marsh, K., Lengeler, C., Snow, R.W., 1996. Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. Trop. Med. Int. Health 1, 139–146.

Ngom, P., Binka, F.N., Phillips, J.F., Pence, B., Macleod, B., 2001. Demographic surveillance and health equity in sub-Saharan Africa. Health Policy Plan 16, 337–344.

Nobre, A.A., Schmidt, A.M., Lopes, H.F., 2005. Spatio-temporal models for mapping the incidence of malaria in Par. Environmetrics 16, 291–304.

Noor, A.M., Gething, P.W., Alegana, V.A., Patil, A.P., Hay, S.I., Muchiri, E., Juma, E., Snow, R.W., 2009. The risks of malaria infection in Kenya in 2009. BMC Infect. Dis. 9, 180.

Nyarango, P.M., Gebremeskel, T., Mebrahtu, G., Mufunda, J., Abdulmumini, U., Ogbamariam, A., Kosia, A., Gebremichael, A., Gunawardena, D., Ghebrat, Y., Okbaldet, Y., 2006. A steep decline of malaria morbidity and mortality trends in Eritrea between 2000 and 2004: the effect of combination of control methods. Malar. J. 5, 33.

O'Meara, W.P., Bejon, P., Mwangi, T.W., Okiro, E.A., Peshu, N., Snow, R.W., Newton, C.R., Marsh, K., 2008. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. The Lancet 372, 1555–1562.

Oduro, A., Koram, K., Rogers, W., Atuguba, F., Ansah, P., Anyorigiya, T., Ansah, A., Anto, F., Mensah, N., Hodgson, A., Nkrumah, F., 2007. Severe falciparum malaria in young children of the Kassena-Nankana district of northern Ghana. Malaria Journal 6, 96.

Oesterholt, M.J.A.M., Bousema, J.T., Mwerinde, O.K., Harris, C., Lushino, P., Masokoto, A., Mwerinde, H., Mosha, F.W., Drakeley, C.J., 2006. Spatial and temporal variation in malaria transmission in a low endemicity area in northern Tanzania. Malar. J. 5, 98.

Okello, P.E., Van Bortel, W., Byaruhanga, A.M., Correwyn, A., Roelants, P., Talisuna, A., D'Alessandro, U., Coosemans, M., 2006. Variation in malaria transmission intensity in seven sites throughout Uganda. Am. J. Trop. Med. Hyg. 75, 219–225.

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 Med 7, 75.

Oti, S.O., Kyobutungi, C., 2010. Verbal autopsy interpretation: a comparative analysis of the InterVA model versus physician review in determining causes of death in the Nairobi DSS. Population Health Metrics 8, 21.

Owusu-Agyei, S., Awini, E., Anto, F., Mensah-Afful, T., Adjuik, M., Hodgson, A., Afari, E., Binka, F., 2007. Assessing malaria control in the Kassena-Nankana district of northern Ghana through repeated surveys using the RBM tools. Malar. J. 6, 103.

Owusu-Agyei, S., Fryauff, D.J., Chandramohan, D., Koram, K.A., Binka, F.N., Nkrumah, F.K., Utz, G.C., Hoffman, S.L., 2002. Characteristics of severe anemia and its association with malaria in young children of the Kassena-Nankana District of northern Ghana. Am. J. Trop. Med. Hyg. 67, 371–377.

Owusu-Agyei, S., Smith, T., Beck, H., Amenga-Etego, L., Felger, I., 2002. Molecular epidemiology of Plasmodium falciparum infections among asymptomatic inhabitants of a holoendemic malarious area in northern Ghana. Tropical Medicine & International Health 7, 421–428.

Paciorek, C.J., Schervish, M.J., 2006. Spatial Modelling Using a New Class of Nonstationary Covariance Functions. Environmetrics 17, 483–506.

Phillips-Howard, P.A., Nahlen, B.L., Kolczak, M.S., Hightower, A.W., ter Kuile, F.O., Alaii, J.A., Gimnig, J.E., Arudo, J., Vulule, J.M., Odhacha, A., Kachur, S.P., Schoute, E., Rosen, D.H., Sexton, J.D., Oloo, A.J., Hawley, W.A., 2003. Efficacy of permethrin-treated bed nets in the prevention of mortality in young children in an area of high perennial malaria transmission in western Kenya. Am. J. Trop. Med. Hyg. 68, 23–29.

Rajaratnam, J.K., Marcus, J.R., Flaxman, A.D., Wang, H., Levin-Rector, A., Dwyer, L., Costa, M., Lopez, A.D., Murray, C.J.L., 2010. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. Lancet 375, 1988–2008.

Ridout, M., Hinde, J., DeméAtrio, C.G.B., 2001. A Score Test for Testing a Zero-Inflated Poisson Regression Model Against Zero-Inflated Negative Binomial Alternatives. Biometrics 57, 219–223.

Riedel, N., Vounatsou, P., Miller, J.M., Gosoniu, L., Chizema-Kawesha, E., Mukonka, V., Steketee, R.W., 2010. Geographical patterns and predictors of malaria risk in Zambia: Bayesian geostatistical modelling of the 2006 Zambia national malaria indicator survey (ZMIS). Malar. J. 9, 37.

Roca-Feltrer, A., Carneiro, I., Smith, L., Schellenberg, J.R.A., Greenwood, B., Schellenberg, D., 2010. The age patterns of severe malaria syndromes in sub-Saharan Africa across a range of transmission intensities and seasonality settings. Malar. J. 9, 282.

Rogerson, S.J., Carter, R., 2008. Severe Vivax Malaria: Newly Recognised or Rediscovered? PLoS Med 5.

Romagosa, C., Ordi, J., Saute, F., Quintó, L., Machungo, F., Ismail, M.R., Carrilho, C., Osman, N., Alonso, P.L., Menendez, C., 2007. Seasonal variations in maternal mortality in Maputo, Mozambique: the role of malaria. Tropical Medicine & International Health 12, 62–67.

Ross, A., Smith, T., 2006. The effect of malaria transmission intensity on neonatal mortality in endemic areas. Am. J. Trop. Med. Hyg. 75, 74–81.

Rowe, A.K., Rowe, S.Y., Snow, R.W., Korenromp, E.L., Schellenberg, J.R.A., Stein, C., Nahlen, B.L., Bryce, J., Black, R.E., Steketee, R.W., 2006. The Burden of Malaria Mortality Among African Children in the Year 2000. Int. J. Epidemiol. 35, 691–704.

Rue, H., Tjelmeland, H., 2002. Fitting Gaussian Markov Random Fields to Gaussian Fields. Scandinavian Journal of Statistics 29, 31–49.

Rumisha, S., Smith, T., Abdalla, S., Masanja, H., Vounatsou, P., 2012. Modeling heterogneity in malaria transmission using large sparse spatio-temporal entomological data. Spatial and Spatio-temporal Epidemiology (Under review).

Rumisha, S.F., 2013. Modelling the seasonal and spatial variation of malaria transmission in relation to mortality in Africa. (Thesis submitted for PhD, University of Basel).

Rumisha, S.F, Gosoniu, D.G., Kasasa, S., Smith, T., Abdulla, S., Masanja, H., Vounatsou, P., Bayesian modeling of large geostatistical data to estimate seasonal and spatial variation of sporozoite rate. Journal of Applied Statistics (Under review).

Rutebemberwa, E., Kallander, K., Tomson, G., Peterson, S., Pariyo, G., 2009. Determinants of delay in care-seeking for febrile children in eastern Uganda. Trop. Med. Int. Health 14, 472–479.

Ruzicka, L.T., Lopez, A.D., 1990. The use of cause-of-death statistics for health situation assessment: national and international experiences. World Health Stat Q 43, 249–258.

Sacarlal, J., Nhacolo, A.Q., Sigaúque, B., Nhalungo, D.A., Abacassamo, F., Sacoor, C.N., Aide, P., Machevo, S., Nhampossa, T., Macete, E.V., Bassat, Q., David, C., Bardají, A., Letang, E., Saúte, F., Aponte, J.J., Thompson, R., Alonso, P.L., 2009. A 10 year study of the cause of death in children under 15 years in Manhiça, Mozambique. BMC Public Health 9, 67.

Saúte, F., Aponte, J., Almeda, J., Ascaso, C., Vaz, N., Dgedge, M., Alonso, P., 2003. Malaria in southern Mozambique: incidence of clinical malaria in children living in a rural community in Manhiça district. Trans. R. Soc. Trop. Med. Hyg. 97, 655–660.

Schur, N., Gosoniu, L., Raso, G., Utzinger, J., Vounatsou, P., 2011. Modelling the geographical distribution of co-infection risk from single-disease surveys. Statistics in Medicine 30, 1761–1776.

Shaukat, A.M., Breman, J.G., McKenzie, F.E., 2010. Using the entomological inoculation rate to assess the impact of vector control on malaria parasite transmission and elimination. Malar. J. 9, 122.

Shililu, J., Ghebremeskel, T., Mengistu, S., Fekadu, H., Zerom, M., Mbogo, C., Githure, J., Novak, R., Brantly, E., Beier, J.C., 2003. High seasonal variation in entomologic inoculation rates in Eritrea, a semi-arid region of unstable malaria in Africa. Am. J. Trop. Med. Hyg 69, 607–613.

Shililu, J.I., Maier, W.A., Seitz, H.M., Orago, A.S., 1998. Seasonal density, sporozoite rates and entomological inoculation rates of Anopheles gambiae and Anopheles funestus in a high-altitude sugarcane growing zone in Western Kenya. Trop. Med. Int. Health 3, 706–710.

Singer, D.J., Willet, B.J., 1993. It's about Time: Using Discrete-Time survival analysis the duration and time of events. Journal of Educational Statistics 18.

Smith, T A, Leuenberger, R., Lengeler, C., 2001. Child mortality and malaria transmission intensity in Africa. Trends Parasitol. 17, 145–149.

Snow, R.W., Craig, M., Deichmann, U., Marsh, K., 1999. Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population. Bull. World Health Organ. 77, 624–640.

Snow, R.W., Korenromp, E.L., Gouws, E., 2004. Pediatric mortality in Africa: plasmodium falciparum malaria as a cause or risk? Am. J. Trop. Med. Hyg 71, 16–24.

Snow, R.W., Marsh, K., 1995. Will reducing Plasmodium falciparum transmission alter malaria mortality among African children? Parasitology Today 11, 188–190.

Snow, R.W., Marsh, K., 2002. The consequences of reducing transmission of Plasmodium falciparum in Africa. Adv. Parasitol. 52, 235–264.

Snow, R.W., Omumbo, J.A., Lowe, B., Molyneux, C.S., Obiero, J.O., Palmer, A., Weber, M.W., Pinder, M., Nahlen, B., Obonyo, C., Newbold, C., Gupta, S., Marsh, K., 1997. Relation between severe malaria morbidity in children and level of Plasmodium falciparum transmission in Africa. Lancet 349, 1650–1654.

Soares Magalhães, R.J., Biritwum, N.-K., Gyapong, J.O., Brooker, S., Zhang, Y., Blair, L., Fenwick, A., Clements, A.C.A., 2011. Mapping Helminth Co-Infection and Co-Intensity: Geostatistical Prediction in Ghana. PLoS Negl Trop Dis 5, e1200.

Somi, M.F., Butler, J.R.G., Vahid, F., Njau, J.D., Kachur, S.P., Abdulla, S., 2007. Economic burden of malaria in rural Tanzania: variations by socioeconomic status and season. Trop. Med. Int. Health 12, 1139–1147.

Stapleton, D.H., 2009. Historical Perspectives on Malaria: The Rockefeller Antimalaria Strategy in the 20th Century. Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine 76, 468–473.

Steketee, R.W., Campbell, C.C., 2010. Impact of national malaria control scale-up programmes in Africa: magnitude and attribution of effects. Malaria Journal 9, 299.

Steketee, R.W., Nahlen, B.L., Parise, M.E., Menendez, C., 2001. The burden of malaria in pregnancy in malaria-endemic areas. Am. J. Trop. Med. Hyg. 64, 28–35.

Stolwijk, A.M., Straatman, H., Zielhuis, G.A., 1999. Studying seasonality by using sine and cosine functions in regression analysis. J Epidemiol Community Health 53, 235–238.

The malERA Group, 2011. A Research Agenda for Malaria Eradication: Monitoring, Evaluation, and Surveillance. PLoS Med 8, e1000400.

Tollman, S.M., Zwi, A.B., 2000. Health system reform and the role of field sites based upon demographic and health surveillance. Bull. World Health Organ. 78, 125–134.

Venables, W.N., Ripley, B.D., 2003. Modern Applied Statistics with S, 4th ed. 2002. Corr. 2nd printing. ed. Springer.

Vounatsou, P., Raso, G., Tanner, M., N'goran, E.K., Utzinger, J., 2009. Bayesian geostatistical modelling for mapping schistosomiasis transmission. Parasitology 136, 1695–1705.

Vyas, S., Kumaranayake, L., 2006. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan 21, 459–468.

WHO | World Malaria Report 2009 [WWW Document], n.d. URL http://www.who.int/malaria/world_malaria_report_2009/en/index.html (accessed 5.15.11).

WHO Report, 2011. WHO | World Malaria Report 2011 [WWW Document]. WHO. URL http://www.who.int/malaria/world_malaria_report_2011/en/ (accessed 5.30.12).

Wirtz, R.A., Zavala, F., Charoenvit, Y., Campbell, G.H., Burkot, T.R., Schneider, I., Esser, K.M., Beaudoin, R.L., Andre, R.G., 1987. Comparative testing of monoclonal antibodies against Plasmodium falciparum sporozoites for ELISA development. Bull World Health Organ 65, 39–45.

Yamey, G., 2000. African heads of state promise action against malaria. BMJ 320, 1228.

Curriculum vitae

Personal details

Name:	Simon Kasasa
Date and place of birth	28th May 1968/Bujumba-Butambala
Nationality	Ugandan
Official Address	C/O Makerere University School of Public Health P.O Box 7072 Kampala, Uganda skasasa@musph.ac.ug / sdkasasa@yahoo.com

Education background

Institution	Degree	Period
University of Basel, Switzerland	PhD (Epidemiology)	2007- 2012
Case Western Reserve University, Cleveland, Ohio, USA	Masters of Science	1997-1999
Makerere University Kampala, Uganda	Bachelors of Statistics (Hons.)	1991-1994

Work experince

1995- date Biostaistician, School of Public Health, Makerere University College of Health Sciences, Kampala Uganda

Selected pulications

1) Patrick E Ogwang, Jasper O Ogwal, Simon Kasasa, Deogratius Olila, Francis Ejobi David Kabasa and Celestino Obua:- *Artemisia Annua L.* Infusion Consumed Once a Week Reduces Risk of Multiple Episodes of Malaria: A Randomised Trial in a Ugandan Community, **Tropical Journal of Pharmaceutical Research** June 2012; 13 (3): 445-453

2) Fred Nuwaha, Simon Kasasa, Godwill Wana, Elly Muganzi and Elioda Tumwesigye**:-** Effect of home-based HIV counselling and testing on stigma and risky sexual behaviours: serial cross-sectional studies in Uganda, **Journal of the International AIDS Society** 2012, 15:17423

3) David Mukanga, James K Tibenderana, Stefan Peterson, George W Pariyo, Juliet Kiguli, Peter Waiswa,Rebecca Babirye, Godfrey Ojiambo, Simon Kasasa, Franco Pagnoni, Karin Kallander:- Acceptability and utilisation of community health workers using malaria and pneumonia diagnostics for case management of febrile children in Uganda, **Malaria Journal 2012, 11:121**

4) Olivia Namusisi, Juliet N Sekandi, Simon Kasasa, Peter Wasswa, Nicholas T Kamara, Medard Bitekyerezo, Placid Mihayo, Sheba N Gitta, David Mukanga:- Risk factors for non-communicable diseases in rural Uganda: a pilot surveillance project among diabetes patients at a referral hospital clinic. **The Pan African Medical Journal. 2011;10:47**

5) Patrick E. Ogwang, Jasper O. Ogwal, Simon Kasasa, Francis Ejobi, David Kabasa and Celestino Obua:– Use of *Artemisia annua L*. infusion for malaria prevention: Mode of Action and Benefits in a Ugandan community, **British Journal of Pharmaceutical Research**. 2011 Volume 1 (Issue 4): 124-132

6) Elioda Tumwesigye, M.S.;Goodwill Wana B.Sc; Simon Kasasa, M.S; Elly Muganzi, M.D and Fred Nuwaha, Ph.D:- High uptake of Home-Based, District-Wide, HIV Counseling and Testing in Uganda: **AIDS Patient Care and STDs**, Volume 24,Number 11, 2010

7) Lynn Atuyambe, Stella Neema, Erasmus Otolok-Tanga, Gakenia Wamuyu-Maina, Simon Kasasa, and Fred Wabwire-Mangen:- The effects of enhanced access to antiretroviral therapy: a qualitative study of community perceptions in Kampala city, Uganda : **African Health Sciences**. 2008 March; 8(1): 13–19

8) Lewis RF, Braka F, Mbabazi W, Makumbi I, Kasasa S, Nanyunja M:- Exposure of Ugandan health personnel to measles and rubella: evidence of the need for health worker vaccination. **Vaccine.** Volume 24, Issues 47-48, 17 November 2006, Pages 6924-6929

9) Braka F, Nanyunja M, Makumbi I, Mbabazi W, Kasasa S, Lewis RF:- Hepatitis B infection among health workers in Uganda: Evidence of the need for health worker protection .**Vaccine** Volume 24, Issues 47-48, 17 November 2006, Pages 6930-6937

10) William Hsiao, Peter Berman, William Bazeyo, Simon Kasasa, Irumba Atuhurra:- Is social health insurance feasible in Uganda? **Uganda Health Bulletin** Volume 7 Number 3 July 2001

11) Francis N Ssali, Moses R.Kamya, Fred Wabire-Mangen, Simon Kasasa, Moses Joloba, Donna Williams, Roy D. Mugerwa, Jerrold J Ellner and John L Johnson:- A prospective study of Community Acquired Bloodstream Infections Among Febrile Adult Admitted to Mulago Hospital in Kampala , Uganda . **Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology** 1998 Dec 15;19(5):484-9.