ENTOMOLOGICAL RISK FACTORS ASSOCIATED WITH LOW-LEVEL TRANSMISSION IN A PRE-ELIMINATION ZONE IN SOUTHERN ZAMBIA

by

Victoria Kamilar

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science

Baltimore, Maryland

May 2022

Abstract

While the sub-Saharan African region carries a disproportionately high amount of the global malaria burden, many historically endemic areas are moving towards elimination. Achieving elimination in areas with low levels of endemicity requires targeted interventions against the remaining vector species, and there is potential that traditional control measures such as long-lasting insecticide-treated nets (LLIN) and indoor residual spraying (IRS) may not be as effective against these species. By performing a longitudinal cohort analysis in a low transmission area, we seek to determine the association between risk factors and malaria vectors species as it pertains to ongoing transmission, to inform how potential interventions should be targeted. This analysis is focused on data gathered by the Southern and Central Africa International Centers of Excellence for Malaria Research (ICEMR) from 2018 to 2020 as part of the Antoomwe Study in Mapanza, Choma District, Southern Province, Zambia. Fifty-nine households were in enrolled in the study for an average of 18-months. Participating households agreed to entomological trapping and the administrations of surveys for epidemiological data. Centers for Disease Control (CDC) light traps were used for monthly collection of mosquitoes inside and outside of households. Collections were performed indoors (n = 1,113) and outdoors (n = 1,021) with 3,095 mosquitoes collected overall. Mosquitoes were morphologically and molecularly confirmed to species, revealing ten anopheline species with *Anopheles arabiensis* being the most prevalent, representing 36% of anophelines identified. Blood meal analysis was performed in addition to *Plasmodium falciparum* detection by ELISA, with only one parasite-positive specimen detected. A mixed effect negative binomial regression was used to determine the association of

known risk factors and control measures with mosquito prevalence. Additionally, spatial analysis was performed to determine household proximity to various environmental features that affect mosquito species composition, of which tree density around the home was significantly associated with increase mosquito abundance. Analysis of residual spatial dependency in model performance across time was also considered. This study aims to contribute to a better understanding of the changing dynamics of malaria vectors in Choma District, Zambia to mitigate malaria transmission.

Advisor: Timothy Shields

Primary Reader: Douglas Norris

Acknowledgments

I thank Dr. Douglas Norris for his wealth of insights and kind words, Tim Shields for being my thesis advisor, and Drs. Michael Desjardins and Frank Curriero on their spatial and statistical guidance. I would like to express gratitude towards Limonty Simubali, Twig Mudenda, and Harry Hamapumbu from MRT for their help with data collection, coordination, and cleaning. In addition, I want to recognize Jessica Schue for her work in the study design, coordination, and utilization of data as well as Mary Gebhardt for her help in these matters. I also appreciate the field mosquito collectors and the participants of the study. Research funding was provided from the National Institutes of Health International Centers of Excellence for Malaria Research (U19AI089680).

Contents

Abstract ii

Acknowledgements iv

List of Tables vi

List of Figures vii

1. Introduction 1

1.1 Transmission Risk 1
1.2 Malaria History of Impact 1
1.3 Zambian Malaria Challenges and Context 3
1.4 Risk Factors to Transmission 5
1.5 Interventions and Vector Control 6
1.6 Vector Composition 8
1.7 Environmental Factors 11
1.8 Measures to Describe Vectors 12
1.9 Issue of Resistance 14
1.10 Control and Interventions 16
1.11 Study Background and Problem Statement 18

2. Methods 20

2.1 Study Design 20
2.2 Ethics and Approvals 23
2.3 Spatial Data 24
2.4 Entomological Data 25
2.5 Analysis 26

3. Results 31

4. Discussion 44

4.1 Overview 44
4.2 Entomological Inferences 45
4.3 Model Inferences 46
4.4 Limitations 48
4.5 Future Work 49

5. Conclusion 50

References 51

Appendix 59

Tables 59
Figures 70

List of Tables

3.1 Baseline characteristics of household collections by trap location 36

3.2 GLMMNB main effect model 40

List of Figures

2.1 Map of household enrollment and cluster designation. 22

2.2 Number of female anophelines by species per month 32

2.3 Mosquito abundance across Mapanza, Choma District, Zambia separated by competence as a malaria vector 33

2.4 Land use map of Mapanza, Choma District, Zambia featuring contributing households 43

Introduction

Transmission Risk

Malaria is a vector-borne disease that is spread through the bite of a female anopheline mosquito carrying a *Plasmodium* parasite. Since malaria is a mosquito-borne disease, risks associated are those impacting the vectorial capacity of the mosquito as well as the development of the parasite. Factors known to affect mosquito density and species composition and as a result, malaria transmission, include environment, human behavior, intervention coverage, housing structure, animal ownership, and vertebrate crowding in a given area.[1,2] Environmental factors extend to climate and proximity to breeding sites. Human behaviors include time spent outside at certain parts of the day or indoors unprotected. Intervention coverage pertains to uptake, access, and adherence to implementing methods such as insecticide treated nets (ITN), indoor residual spraying (IRS), and anti-malarial therapies. Housing structural factors that may affect mosquito presence near humans include wall type, open eaves, and sealable windows and doors.

Malaria History of Impact

Malaria is present in many regions across the globe. Areas with the highest concentration of cases are in 10 sub-Saharan African countries earning the attention of global leaders. Malaria is a major cause of global mortality, resulting in more than 409 thousand deaths annually with a disproportionate amount of those cases occurring among children under the age of five years old.[3,4] Over the last few decades, major efforts have been made to improve and implement malaria control programs for the ultimate goal of local elimination followed by global eradication. The World Health Organization (WHO)

and Roll Back Malaria (RBM) Partnership has called international attention to the countries of greatest impact in an approach titled 'High burden to High Impact' which describes a focusing of intervention efforts in the areas with the highest burden of malarial disease.[5] This comes after the resolution to reduce malaria by 75% from levels seen in year 2000 by the year 2015 failed to reach its intended target.[5] Significant reductions in morbidity and mortality have been observed since the Millennium Development Goals were set forth by the WHO in 2000, but a leveling off of effectiveness has driven the need for new approaches.[5] In 2015, the World Health Assembly promoted The Global technical strategy for malaria 2016-2030 (GTS) which proposed the new target of a 90% reduction in malaria incidence and mortality rates globally by 2030 as well as including the more incremental goal of a 40% reduction in incidence by 2020.[3] This goal was not achieved, but in many countries the initiatives were useful in decreasing incidence substantially. [3] Recommendations include a shift in focus from using morbidity and mortality as sole measures of efficacy of interventions to focusing on coverage of interventions as well as other factors allowing for efficient transmission to better approach the remaining disease burden.[3] While these approaches led to numbers trending in the proper direction, with the onset of the COVID-19 pandemic many services were paused which resulted in a global increase in incidence and mortality between 2019 and 2020, 95% of which was contributed to the WHO African region.[6] This reversal illustrates the impact of vector control and health care interventions and services on transmission. This also implies that until global eradication is reached interventions cannot be relaxed or resurgences will occur.

Zambian Malaria Challenges and Context

Zambia, a lower-middle income landlocked country in south-central Africa is home to 18.4 million people and 3.4 million cases of malaria annually.[6] Zambia also contains a heterogenous landscape of malaria transmission making it a prime location for the study of intervention efficacy and transmission dynamics. The Southern and Central International Center of Excellence for Malaria Research (ICEMR) works out of two study regions in Zambia, one in Nchelenge District in the northern wetland portion of country and the other in Choma District in the drier southern region with the help and collaboration of local research groups, government, and communities. The Zambian government has run the National Malaria Elimination Program for decades with collaborative effort from the U.S. with the Presidents Malaria initiative (PMI) which started in 2005, the WHO with the Roll Back Malaria (RBM) Campaign, as well as the Malaria Operation Plan.[7,8] As a result, Zambia has seen success over the last twenty years in ramping up control measures and earning its place as a candidate for elimination in the near future as one of the E8 countries with the goal of elimination by 2030.[9] Interventions include the initiation of using intermittent preventive treatment in pregnancy (IPTp) and artemisinin based combination therapy (ACT) as standards of care for uncomplicated malaria since 2003 and a large scale Integrated Vector Management (IVM) program with initiative to provide indoor residual spraying (IRS), larviciding, environmental management, and ITN distribution since 2004.[1] Gaps in funding, unequitable allocation of resources, and increasing mosquito resistance to insecticides have slowed progress as well as the challenges associated with providing adequate coverage of interventions in areas of the country with poor infrastructure.[9] As of 2020 Zambia comprises 1.4% of

global malaria cases and deaths.[3] Coverage of interventions based on 2016-2018 surveillance data indicated suboptimal access and use of many of these interventions as well as the severe effect the COVID-19 pandemic had on service offerings in 2020.[3] This data indicated only 25% of the intended ITNs for distribution made it into the community and there was a 30% decrease in malaria testing.[3] Even so, the proportion of probable cases confirmed by RDT has been steadily increasing each year since 2014 and between 2019 and 2020 no statistically significant increases in mortality rate have been observed.[3]

Zambia, being heterogenous in malaria case intensity means it is comprised of both high and low transmission areas, classifying the country as low to moderate transmission overall. High transmission areas are characterized by intense vector-human interaction and a parasite prevalence of 50% or greater as defined by NIAID or an Annual Parasite Incidence (API) of greater than 450 per 1000 according to WHO classifications.[10,11,12] Moderate transmission has an API falling within 250-450 per 1000 and low transmission is when malaria is still circulating at low endemic levels with occasional outbreaks. This shift in transmission intensity is marked by a reduction in vector-human interaction, a parasite prevalence less than 50%, and API between 100-250 per 1000.[10,11,12] Very low or sometimes referred to as pre-elimination classifications, near the absolute of zero locally acquired infections, with an API of less than 100 but greater than 0 per 1000 API.[10,11,12] A maintained zero is considered elimination but this still requires continued control and surveillance to prevent establishment of transmission from recuring.[13] Transmission is seen on a continuum rather than the previously categorized "control", "consolidation", "pre-elimination", and "elimination" stages that did not consider the importance of integrated and fine-tuned approaches to focal areas rather than

on a country-wide scale.[14] Elimination is still considered on a country wide scale however, and is defined by an excess of three cases acquired in-country of the same parasite species per year for three years in a row.[14] With this definition, to maintain elimination status the main role of surveillance is to quickly detect cases or foci originating in or out of country, to prevent focal transmission with the use of public health interventions. Similarly, in moderate to low transmission countries heterogenous in transmission intensities, such as Zambia, the necessity for finding imported cases can be significant as they have a large impact on future transmission in an area. This is done by considering human movement. Individuals can become infected in a high transmission zone then return to a low transmission zone and serve as a source of parasite leading to sustained transmission in the transplanted area. The majority of human-related movement of parasite is within an endemic country, but between areas of varying transmission intensity.[15] Choma District in the southern province of Zambia is a low transmission area while Nchelenge District in the north of Zambia is a high transmission setting, making information on recent travel of household members and visitors to the household in epidemiological surveillance an important factor to include.[16,17]

Risk Factors to Transmission

Several studies have been conducted to elucidate the risk factors associated with malaria transmission and mosquito burden near and within the home. These models are typically based in high transmission settings, as the data is easier to gather and more statistically powered. Data gathered in high intensity settings to estimate transmission are typically passively detected mortality and morbidity cases reported by the health care

system as well as using all-cause child mortality as a proxy for malaria diagnosis in malaria endemic countries.[18] In lower incidence areas, a more sensitive measure of transmission dynamics is needed. Many programs opt for active case detection and some champion the use of entomological inoculation rates (EIR) as a metric of transmission intensity not unlike malaria incidence, considering these measures have a strong linear relationship, but EIR relies on human biting rates and rate of infection in vector mosquito species.[6,12,18,19] With active case detection, focal transmission can be observed and traced to observe a clustering of cases which is often missed by studies designed with random or probabilistic sampling of participants rather than selection of high impact foci.[18,20] Considering focal transmission can also be useful in determining the role importation of cases from geographic areas outside the region of study play on sustaining transmission as well as if the remaining competent vector species are being adequately targeted by intervention methods.[21]

Interventions and Vector Control

From 2000 to 2017, reductions in transmission globally were due to three main interventions: distribution of ITNs, IRS, and increased focus on fast case detection and treatment.[5] Each aims to interrupt transmission at a specific point in the vector-human interface. ITN or LLINs work by impregnating a fibrous material with a binding agent and insecticide, typically a synthetic pyrethroid, for up to four years of efficacy given proper care.[22,23] These nets act as a repellent as well as an adulticide if the mosquito makes contact with the net, in addition to a physical barrier. This intervention acts to protect the user of the net and indirectly, others in the room from mosquito bites indoors

during the night. IRS application is a periodical treatment of the indoor walls of a home with insecticide, ideally this is done with a rotation of two to three insecticides belonging to different classes or modes of action.[24] IRS can be expensive to implement and is difficult to do so evenly, considering the variation in supportive infrastructure. IRS also works by providing spatial repellency and contact toxicity which targets endophilic mosquitoes that rest within the home after feeding, resulting in a reduction in daily survival rates of vector species.[23] Case detection and treatment is important to intervene and prevent or reduce uptake of *Plasmodium* gametocytes, the infectious stage of the parasite, to a competent vector species which can then lead to development and amplification within the mosquito that can be transmitted to other humans in the community.[25,26] In Zambia, the first line of treatment for uncomplicated malaria cases is artemether-lumefantrine (AL) or artesunate (AS) for severe malaria.[3] Many African countries use combination therapies such as artemisinin based combination therapy (ACT) for their primary treatment.[11] Most effective of these intervention strategies alone is net usage, but this is likely due to the higher distribution and accessibility associated.[7] LLINs are considered the most cost effective and widely distributed methods of malaria intervention but even so, the WHO recommends behavior intervention campaigns, one of which is titled the "Hang-up Campaign" to promote continued and proper usage of nets over time.[7] Amongst other countries, Zambia adapted this strategy into its malaria operational plan to inform, educate, and foster communication that will lead to increased adherence to the proper use of this intervention.[7] Several factors have been observed to be associated with individual net usage including seasonality, relative distance to a health center, and presence of total mosquitoes including non-vector species.[22] While net usage

is effective and efficient, integrated vector management comprising all of the above interventions as well as being contained within a multifaceted collaboration between the health and public sector is considered the most successful approach.[3,18]

IVM is comprised of epidemiological as well as entomological evaluations and surveillance to inform decisions on how to better shape approaches and target factors of highest importance in the implementation of vector-targeted interventions. Past work in low transmission areas of Zambia have indicated a need for more examinations of environmental features and has suggested a switch from the commonly associated vector species to more cryptic species such as *Anopheles squamosus* which may have different bionomics, reducing the effectiveness of standard malaria preventatives.[11] Bionomics describe the behaviors, activity, seasonal abundance, and ecological aspects of particular species of mosquitoes, which are relevant to tailored control approaches.[27] Behaviors of interest include preferences of host species, foraging location and time, environment features of inhabitance and resting areas, and features of oviposition sites. Prior studies have found complex patterns of efficacy with IRS pertaining to vector bionomics with the specific chemical formulation and the frequency of treatments playing a role.[11] This impact on control is particularly pronounced in areas of low transmission that aim for elimination, such as Choma District. Additionally, evaluating efficacy is difficult with limited data making well-constructed longitudinal studies particularly important to capture effects over time and their association with other changing factors.

Vector Composition

At the Mapanza field site in Choma District, Zambia, the primary vector is *An. arabiensis* a member of the *An. gambiae* complex. *Anopheles arabiensis* is considered to have a higher degree of behavioral plasticity than some other vector species making it more capable of circumventing the protective effects afforded by the primary vector control interventions used in the region.[1] While *An. arabiensis* foraging is variable, prior work in Choma District suggests peak biting times around ten pm or as late as midnight, while in some areas they are considered relatively crepuscular, exophagic, and highly anthropophilic.[1,28] With behavioral plasticity comes added challenges to control as mosquitoes that have adapted to crepuscular feeding, as compared to night feeding, can be well suited to evasion of indoor and sleep centric interventions such as ITN and IRS, as feeding often occurs outdoors in the evening when humans gather for community events or in the mornings when individuals must leave the dwelling early to begin work. Understudied vector species present in the study area include *An. rufipes, An. maculipalpis,* and *An. pretoriensis* which are zoophilic but capable vectors of *Plasmodium* and have been collected and positively tested for sporozoites in prior studies.[28,29] *Anopheles pharoensis* and *An. squamosus* have also been found to be *Plasmodium*-positive but are heavily zoophilic, and thus they have not historically been considered vectors of concern in Zambia.[28] They are classified as major and secondary vectors in Cameroon and Senegal, however.[28] *Anopheles funestus* complex and *An. gambiae* complex are both variable in host preference and vector competence.[28] While *An. funestus* sensu stricto is considered a main vector on the African continent, other species in the complex such as *An. rivulorum* and *An. vaneedeni* are considered largely zoophilic, although occasionally parasite positive earning their distinction as secondary

vectors.[28] *Anopheles leesoni,* also within the *An. funestus* complex, is considered to not play a major role in transmission as their strong zoophilic behavior has historically resulted in low human-contact, although *Plasmodium*-positive specimens have been reported.[28,29] *Anopheles longipalpis* on the other hand, while zoophilic, in Choma District has been reported as highly endophilic and morphological similarity to *An. funestus* make this species easily confused with important vector species.[28]

To determine species identity, morphological identification based on dichotomous keys is typically used to sort specimens into major categories by genus and possibly to species, but often it can be impossible to accurately identify to species or distinguish between known species and new cryptic species by morphology alone.[2,28] Molecular techniques such as polymerase chain reaction (PCR) targeting the internal transcribed spacer 2 (ITS2), mitochondrial 12S, cytochrome oxidase 1 subunit (COI), or other genes or gene fragments have been discovered to aid in distinguishing between species of mosquitoes.[30] Sequencing, ranging from DNA fragments to whole genomes, is often used to compare within databases such as NCBI's GenBank for verification of specimen identifications or to confirm the discovery of new species.[30]

In habitats capable of supporting anopheline life cycles, there are typically other subfamilies and genera of mosquitoes present as well. Culicine mosquitoes, capable of transmitting human diseases including lymphatic filariasis, Chikungunya, and West Nile are often regarded as nuisance mosquitoes in many malaria endemic countries.[31] While *Culex* species are known to exist in great abundance and act as a vector for several pathogens they are typically not heavily investigated in areas of malaria endemicity. More often when they are collected in malaria-geared studies the specimens are

discarded, or if enumerated, are not molecularly confirmed to species.[7,32] If *Culex* data is included in publications at all, they are reported as a quantity to show the effect of nuisance biting on human behavior and adherence to malaria interventions.[7,32] There is suggestive evidence that culicine abundances are affected by malaria vector control.[31] There is also concern that not targeting these species directly will lead to increased insecticide resistance in culicine mosquitoes that will make future work to control the pathogens they transmit more difficult.[33] This concern is not unfounded as *Culex quinquefasciatus* is widely resistant to pyrethroids used in ITNs globally.[34]

Environmental Factors

The lifecycle of the mosquito promotes seasonal variation in relative abundance with numbers of adults often peaking in the rainy season.[35,36] In southern Zambia, the rainy season is from November to April, while the cool-dry season is from May to August and the hot-dry season from August to October which are sometimes combined to be referred to as the dry season.[8] The rainy season has been correlated with higher malaria transmission intensities in prior studies.[8] This correlation can be explained by the effects of climate on the lifecycle of the mosquito. Rain provides breeding sites and greater humidity which can increase the rate of intergenerational development as mosquitoes are found to progress through life stages to adulthood faster in warmer environments.[37] Breeding sites suitable for anopheline mosquitoes to lay eggs in and for larvae to develop are still water.[38] Conventionally, anophelines were predominantly limited to clear water such as edges of lakes, ponds, marshes, and rice paddies, or ephemeral pools after rain, but urban malaria studies have found breeding habitat suitable

to support the anopheline life cycle to include but not be limited to broken water pipes, shallow pooling in agricultural or construction sites, and catch pits.[38] Humidity also increases daily survival as anophelines are not resistant to desiccation, but there are differences in susceptibility to drought and aridity. For example, across the *An. gambiae* complex, *An. arabiensis* is generally considered more drought tolerant than *An. gambiae.*[37,39] Climate as well as entomological data can be used more effectively on a fine or high-resolution scale since many countries are heterogenous in terms of climatic zones due to a variety of factors including elevation, proximity to bodies of water, and land use.[37] In Nchelenge District, Zambia, *An. funestus* abundance peaks in the dry season contrary to the dynamic of most vector species, but in this instance it is suggested to be driven by improvements to habitat survivability as this is the time of year when the marshes are not continuously washed out by rain and the water becomes relatively still and clear.[40] In a study from southern Cameroon, unexpectedly higher overall vector abundance was observed in the dry season but this was attributed to agricultural practices that alter habitat suitability beyond the climate of the region.[41] Microclimates are, on a very fine scale, describing the environment immediately surrounding a mosquito habitat. Microclimate data has shown variation in the rates of development from egg to larvae to pupae by water temperature immediately surrounding the immature mosquitoes as well as the proclivity of endophilic or exophilic activity by ambient temperature indoors compared to those directly around the home.[42]

Measures to Describe Vectors

The effect of climate reaches beyond the development and survivability of the mosquito to the development of the parasite as well.[42,43] Modeling work has suggested that the optimal temperature for malaria transmission is between 25 and 28 degrees Celsius.[42,44] The extrinsic incubation period (EIP) describes the time period of sporogony, defined as the duration of parasite development within the mosquito from infectious blood meal to infectious stages present in the salivary glands.[45] This metric is a temperature dependent, although not exclusively, component of transmission that can be used to estimate the basic reproduction number since it has influence over the capability of mosquitoes to become infectious.[45] EIP are typically shorter duration in higher temperature regions compared to lower temperature regions, but the extent these incubation time frames depend on ambient temperature is determined on a spatial scale as well as a temporal scale since fluctuations have been shown to attenuate the effect extreme temperatures have on developmental duration.[42,46] This is further evidenced by observable variations in EIP between endophilic and exophilic vectors.[42,47]

In addition to EIP, vectorial capacity (VC) is a classic metric used to conceptualize transmission dynamics and intensity with a variety of changing factors. Another metric, relating to VC and EIP is EIR or entomological inoculation rate. EIR is very similar to VC except it describes the actual number of infectious bites rather than the potential number of infectious bites given ideal conditions for infectiousness.[48] Considering these parameters and formulas are useful as ways to conceptualize the effect of interventions on transmission with the goal of targeting each point of vulnerability. In order to perform these analysis, however, a number of variables often not ascertainable in the field must be collected incorporated, this being especially infeasible for

entomological collections taking place within the context of a broader epidemiological study, thus limiting the useful of these metrics. In addition, in areas of low transmission, collecting enough parasite-positive mosquitoes to have power behind EIR calculations can be challenging.

Issue of Resistance

Areas with moderate success in vector control are often comprised of a different or variable set of mosquito vectors that are differentially affected by interventions, resulting in a need for modified approaches to target the behaviors and feeding patterns of the remaining mosquitoes and parasites.[2] Poorly planned control measures have the potential to create larger problems that reach beyond the targeted areas.[49] If control is not improved, the probability of arthropod adaptation will continue to contribute to the growing concerns surrounding resistance.[49] Resistance to control strategies can exist within the mosquito vector as well as the *Plasmodium* parasite and is often considered an inevitability in large scale vector or parasite control programs regardless of effective application.[50] On the African continent the primary parasite causing human malaria is *P. falciparum* which is globally the most fatal of the five known species to cause human malaria.[21] During the Global Malaria Eradication Campaign (GMEC), chloroquine was the primary antimalarial used, but after widespread resistance developed the program appeared doomed to fail and ultimately ended in the late 1960s.[51] Currently, the first line of treatment for *P. falciparum* in many African countries is artemisinin-based combination therapies, artemether-lumefantrine (AL) being the primary artemisinin-derived medication used in Zambia for uncomplicated malaria.[3] Studies focusing on

PfKelch13 and R6221 as markers of adaptations in the parasite that confer drug resistance to these drugs are increasing in prevalence, particularly in southern Africa.[3] Monitoring of drug resistance to human malaria is of great importance to the capability of future malaria treatment and thus therapeutic efficacy studies (TES) on all first and second line antimalarials have been incorporated into the WHOs recommended routine practices.[52] Insecticide resistance, when the mode of action is no longer effective against the mosquito is also growing with the long-standing use of chemical interventions. DDT and pyrethroid resistance is common across several species of malaria vectors in over 40 malaria endemic countries, while sensitivity to carbamates and organophosphate insecticides remains more intact at this time.[3,53] Pyrethroid resistance is a major threat as this is the class of insecticides used almost exclusively in ITN's which are the most cost effective and wide-scale coverage intervention, although in the absence of effective insecticide, nets do confer some protection simply via physical barrier.[50] There is also the impact of behavioral resistance, where interventions have selected for populations of mosquitoes exhibiting behaviors that allow them to be less effectively targeted; behaviors such as diurnal, crepuscular, and outdoor feeding.[50] The Global Plan for Insecticide Resistance Management in malaria vectors (GPIRM) called for monitoring of both physiological and behavioral insecticide resistance to be a part of routine entomological surveillance in 2012 and has since been updated to meet logistic constrains in implementation.[54] The approach to dealing with resistance in vectors is outlined by five pillars and split into short, medium and long term goals.[54] These pillars are "Plan and implement insecticide resistance management strategies in malaria-endemic countries", "Ensure proper, timely entomological and resistance monitoring and effective data

management", "Develop new, innovate vector control tools", "Fill gaps in knowledge on mechanisms of insecticide resistance and the impact of current insecticides resistance management strategies", and "Ensure that enabling mechanisms (advocacy, human and financial resources) are in place".[54] This approach draws on lessons learned as many have stated the Global Malaria Eradication Programme's (GMEP) top-down approach to eradication campaigns failed in the 1950s and 60s as the countries did not have the local public health infrastructure to sustain interventions after funding from external sources lagged due to the demoralizing effects of chloroquine and DDT resistance.[9] This approach also acknowledged that well timed strategies are important, as lapses in effective control can result in increases to pathogen transmission as demonstrated by global trends during the temporary cessation of services following COVID-19 related interruptions in normal services.[55] It is paramount for evidence-based decision making to lead continual integrated intervention programs that consider the context of susceptibility.

Control and Interventions

Considerations of effective surveillance and as a result, interventions, takes into account the human epidemiological factors affecting transmission. Risk factors commonly found to have an association with transmission include housing structure, breeding sites around the home, access to healthcare, the use of nets, IRS, smoke, time spent outdoors during peak biting hours, sleeping outdoors, and animal ownership. Qualities of housing structure that can affect transmission are wall and roof materials, the presence of eaves, and screening on windows and doors with a proper seal.[56,57] Homes with metal roofs, closed eaves, and screened doors are labeled modern homes and show

lower abundances of mosquitoes which reduces transmission.[57] Although, if not well-ventilated temperatures can be very high causing poor adherence to behavioral interventions such as sleeping in a sealed interior, resulting in an attenuation of the effects on transmission.[57] Comfort, convenience, coverage, and cost are important considerations to an effective risk reducing strategy. Access to healthcare as a factor in malaria transmission pertains to the ability to seek diagnosis, treatment, and prenatal care (IPTp). This is considered one of the core components of the RBM action plan.[55] Points of concern are system-wide availabilities of services by location, care-seeking behaviors of the individual, and barriers to service via the distance to and quality of infrastructure between homes and centers where services exist.[58] Regardless of the reason, inadequate access to care increases the probability of serious adverse events, of underreporting of case incidences, and increases duration of infectivity all resulting in a poor prognosis for malaria elimination efforts. Use of interventions such as nets and IRS depend largely on the local control programs funding, coverage, and focus on information campaigns particularly in the case of nets. Net usage depends heavily on the perception of risk.[7,24] Smoke as a primary mosquito preventative, time spent outdoors at peak biting hours, and sleeping outdoors also depend heavily on information campaigns and local cultural practices and perceptions.[59] Time spent away from the home during funerals, religious events, and visits to friends or family can present social barriers to preventative-use as the use of spatial repellants and ITN can be met with social disapproval.[59] The increased likelihood of acquiring infectious bites during high risks activities such as these are compounded by the fact that they are often done away from the home, increasing the chances for transplantation of parasites to potentially lower intensity settings around the

home. Animal ownership and proximity to breeding sites depends on the needs of the individuals in households and are often not modifiable by intervention. These factors are typically included in information collected to estimate risk and where other interventions should be implemented. Animal ownership modulates transmission dynamics as animals can be a chemosensory attractant as well as a host capable of sustaining mosquito populations away from insecticides while still allowing for the risk of transmission to humans to be high in areas around the home or place of work.

Consideration of human epidemiological factors is important to the design and implementation of control programs. Surveillance of healthcare data, climate, and entomological population dynamics are not sufficient without data on intervention acceptability, coverage, and the consideration of factors that may vary the effectiveness of control practices. To gather human epidemiological data, surveys and meticulously designed cohort studies need to be implemented. Mobility pattern data has been collected using cellphone data but this still misses a large portion of fine scale travel.[60] Healthcare data misses the majority of asymptomatic or subclinical infections involved in sustained transmission.[18] Climate data on a fine enough scale to inform changes to transmission are typically only collected in research studies, rather than as a component of routine surveillance. Most often this type of data is inferred from remotely sensed data for larger study areas or for programmatic uses. Therefore, integrated approaches that rely on the public sector, healthcare, community engagement, and research collaboratives are needed to provide adequate data to inform governmental and funding decision making.

Study Background and Problem Statement

The Southern and Central African International Center of Excellence in Malaria Research (ICEMR) work collaboratively with the Macha Research Trust (MRT) in Macha, Choma District, Southern Province of Zambia. The catchment area for the Antoomwe Project is a rural area comprised of villages that have homesteads of which are single or multi structure and often house an extended family. As a single unit each collection of structures are referred to as households. The ICEMR, established in 2010, began work in southern and central Africa with four field sites, two in Zambia, one in Zimbabwe, and one in DRC investigating topics ranging from transmission dynamics and efficacy of intervention programs to entomological factors and insecticide resistance across variations in environment types and transmission intensities. The MRT was established in 2003 as an independent research entity with recognition by the Zambian government.[61] The collaboration between Johns Hopkins affiliates and MRT extends to the early 2000's and has allowed for the implementation of numerous studies including those within Choma District, Zambia.[61]

The Antoomwe project was nested within larger ICEMR projects that consider country-wide malaria transmission dynamics and reactive test-and-treat programs with the goal of refining reactive strategies and vector control programs in countries moving towards malaria elimination. This thesis builds on data collected from programmatic household surveys and entomological collections as well as continually monitored environmental data for analytical techniques to be leveraged to make progress towards this goal. This thesis aims to address the following objectives:

- Identify household level factors associated with mosquito abundance as a measure of malaria transmission risk in Mapanza, Choma District, Zambia using a multilevel modeling framework.
- Identify associations between vector abundance and human behaviors relevant to transmission risk in Mapanza, Choma District, Zambia.
- Determine strength of association between fine scale environmental factors and vector abundance using spatial analysis.

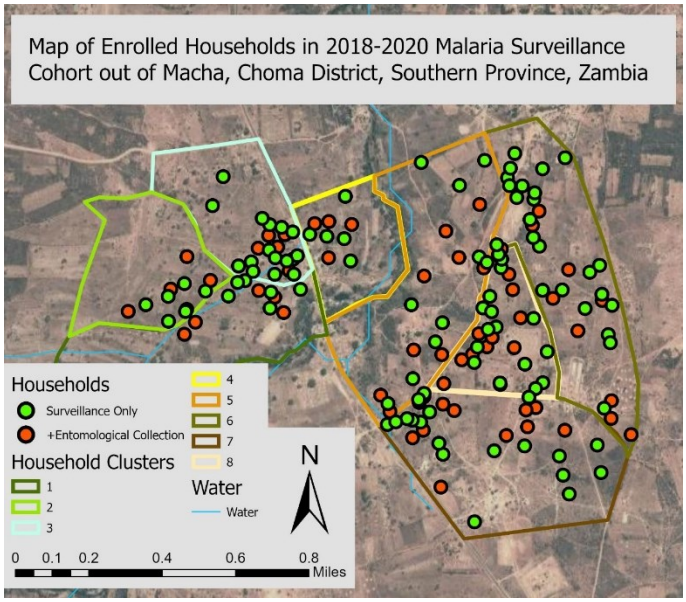
Understanding how risk factors are modified by transmission setting is important to designing targeted and efficient approaches to achieve sustainable local elimination. Determining the association between household factors and vector species abundance and composition will allow for increased information on the heterogeneity of risks for malaria resurgence in areas of low transmission during routine national IVM implementation. Identifying links between vector abundance and human behaviors as they relate to intervention use and adherence to low-risk practices will allow for improved community engagement efforts to broach topics of the importance of maintenance despite low transmission in pre-elimination settings. Considering the granularity of spatial dependance visible in the association between vector abundance and environmental factors in a small catchment area such as the Antoomwe cohort within Mapanza, Choma District of southern Zambia will provide information useful to future surveillance efforts that aim to capture and characterize associations on the micro-epidemiological scale.

Methods

Study Design

The Antoomwe study focuses on a catchment area located in Mapanza, Choma District of Southern Province, Zambia and represents a low malaria transmission area. The study was designed as an open-enrollment prospective longitudinal cohort of geographically contiguous households arranged into 6 clusters and of which a random sampling of 59 households out of the 202 households surveilled were included in monthly entomological collections, while all households provided monthly-repeated epidemiological surveillance data (Figure 1). Longitudinal collections of epidemiological and entomological data was included in order to elucidate changes in vector species and human behaviors with changes in climate, vector-control intervention use, and time. Hourly climate data is collected from the HOBO field station at the MRT field station. This data is aggregated to day (6am to 6pm) and night averages for each month. Households were selected through digitization of high-resolution satellite imagery, after which point the field team visited each household to receive informed consent from the head of house for enrollment into the study. Each household was given a unique identification number for the duration of follow-up. In the event of a change in occupancy the new owners retain the household identification number but if individuals move to a new household, that new household will be enrolled with a new identification number. Participants who provided information via survey response or biospecimen collection were also provided a unique identifier. Each individual participant was provided a study identification card with their unique participant identification, cluster number, household identification and enrollment date listed on the card.

Figure 1. Map of household enrollment and cluster designation.



At the first visit, the epidemiological data collected included participant demographic characteristics, malaria symptoms and treatment history, travel history, health seeking behavior, the use of ITN over specified periods of time and how many ITN are used in the household, time spent outdoors, and socioeconomic information. Upon each visit within the follow up a period, the survey questionnaires administered by the field team aimed to evaluate time-varying factors such as interim illnesses, travel, and net usage. Household surveys were conducted to gather information on the household construction, animal ownership, transportation accessible, household amenities, and the receipt of malaria interventions such as IRS, mass drug administration (MDA), or reactive test and treat care. Entomological collections were accompanied by surveys collecting information on relevant data that recorded the location of the mosquito trap, including structure type, construction, occupancy, animal presence, and use of vector control measures. Information was collected for up to 24 months of follow up time per household, after which all households were administratively censored. Surveys were

administered by trained interviewers and participants could decline to respond to any question for any reason.

Ethics and approvals

Consent forms were collected for all adults participating in the study as well as parental permission and child assent forms for those under the age of 18 who participated in the study. All forms were developed in English by the JHBSPH study team and translated into Chitonga, the local language, by the MRT team. For entomological field data collection, oral consent by an adult assuming responsibility for the household was obtained at each visit before any traps were set. The protocol detailing the study design, implementation specifics, and data collection tools and consent forms was approved by both the Zambian Tropical Diseases Research Center Ethics Review Committee in April 2018 and the Johns Hopkins Institutional Review Board (IRB) , as an amendment to the original ICEMR study protocol approval. After all ethical approvals were completed, ICEMR and MRT approached the provincial and district health directors to obtain approval to work in their communities. Starting June 2018 community outreach began to involve the community leaders in discussion and devise a community advisory board. Community leader meetings provided an opportunity to describe the intended study, allow for questions to be answered and have the ability to circulate printed summaries of the study. The community advisory board was comprised of 10 members and included village heads and Mapanza health center staff to aid in the representation of community considerations as well as assist in the dissemination of interim study results to the people over the course of the study period.

Spatial Data

In order to initially enumerate and select households for inclusion into the study, high-resolution Worldview satellite images were obtained from Apollo Mapping (Boulder, CO) and were digitized. After enrollment into the study, households were designated into one of nine contiguous clusters. Finalized household location data was collected through the use of GPS enabled tablets and was recorded at each household during the follow up period. In the event of tablet malfunction, GPS devices were used and coordinates were entered into datasheets manually. Coordinate locations were inputted into ArcGIS (ESRI, Redlands) and converted to the UTM zone 35 South coordinate system.

Location data was used to find associations between environmental features and household vector populations as well as to describe residual spatial dependency in the candidate models. Environmental features were extracted in ArcGIS through supervised classification of an adjusted orthomosaic color balance from 2-meter resolution multispectral satellite imagery. By leveraging the spectral bands, contrasts in land type was used to create a classification scheme to distinguish between landcover and vegetation, specifically water, and vegetation such as grasses, shrubbery, and trees around structures. The ISO Cluster Unsupervised Classification tool in ArcGIS was used to classify land type, then manual classifications were made to train the classification tool to make a more accurate classification scheme. Buffer zones with a radius of 50 meters surrounding each household were created to calculate the percent land type immediately around each household. Distance from household coordinates to the nearest bodies of

contiguous water was calculated. Bodies of water include streams, rivers, and any land type classified as water that is greater than 20 square meters in area. This definition is used to reduce the risk of misclassification. The data generated was then able to be incorporated in statistical modeling of mosquito abundance as a function of contextual risk factors.

Entomological Data

At each household included in the entomological collection, CDC miniature light traps were set indoors or outdoors. Indoor traps were set at the head-end of the sleeping area 1.5 meters above the ground. Outdoor traps were set at the same height next to a cooking area or near an animal pen or enclosure. Traps were not chemically baited, the proximity to humans or animals was used as the mosquito lure. The light and fan components of the trap were turned on at 18:00 at all sites and turned off at 6:00 the following morning. Mosquitoes collected were then transported to the MRT laboratory, killed by freezing, and processed for morphological identification to sort by subfamily into culicine, anopheline, or other. For the anophelines, morphologically identification was extended to classifying specimen into probable species to then apply the proper molecular test to confirm species identifications. Proper morphological identification was promoted by interagency training workshops between the Johns Hopkins ICEMR and MRT team and followed taxonomy laid out in a dichotomous key on Afrotropical anopheline mosquitoes.[62]

After morphological ID, specimen were characterized as visually blooded or not visually blooded. Mosquitoes were then split between the thorax and abdomen. Both

sections of the mosquito were homogenized and DNA was extracted using the Marriott Mosquito Extraction Protocol. [63] Anophelines were identified to species by PCR, targeting the ITS2 region of nuclear rDNA to differentiate between amplicons of varying sizes depending on species when run using gel electrophoresis.[64] The ITS2 PCR does not differentiate among members of the *An. gambiae* and *An. funestus* complexes, so aliquots of the specimen that belonged to those groups, with expected band sizes being 850 basepairs (bp) for *An. funestus* and 600 bp for *An. gambiae*, were then run using assays designed to do so. The minimized *An. gambiae* complex PCR assay contains four primers that are designed to produce three differently sized amplicons of the rDNA spacer region which will show *An. gambiae* s.s. at 390 bp, *An. arabiensis* at 315 bp, *An. quadriannulatus* at 150 bp after gel electrophoresis and imaging.[65] The *An. funestus* complex PCR assay targets the ITS2 regions of nuclear rDNA with seven species specific primers and one forward primer which will show expected product sizes at 505 bp for *An. funestus* s.s., 146bp *for An. leesoni*, 587bp for *An. vaneedeni*, 252bp for *An. parensis*, 313bp for *An. rivulorum-like*, and 390 for. *An. funestus-like*.[65] Positive and negative controls were obtained from insectary specimen of known species. ELISA analysis was used to detect the circumsporozoite protein (CSP) of *P. falciparum* in female anophelines.[66]

Analysis

The dataset used for analysis was created through merging forms of different types, as defined by the household, trap type, and date information. Data cleaning was performed through an iterative process of logical checks and formalizing data queries

with the field team to develop corrections. Data with queries that could not be resolved were omitted from analysis. Exploratory data analysis included defining variable types as well as making figures and tabulations to clarify relationships and correlations between variables. Before model construction, conceptual frameworks were built based on past research and modified by the completeness of the dataset. There were three frameworks developed prior to analysis with covariates focusing on environmental factors, location data, human/animal interaction, behaviors, and the use of prevention methods.

Univariate analysis was performed on all risk factors of sufficient response size for all applicable primary model types, given the type of data. Generalized Linear Mixed Effects Negative Binomial Regression Models (GLMMNB) by maximum likelihood (Laplace Approximation) was used to describe the relationship between risk factors and mosquito counts at each household, on a monthly time scale. Seasonality, defined as Rainy (November to April) or Dry (May to October) based on prior classification in the literature, was also considered.[8,20] Household number is included in the random effects of all models to describe within house correlation of data across repeated visits. GLMMNB modeling failed to find differences of >5% between covariate outcomes of univariate models of the same risk factor defined by y = total mosquito count versus y = anopheline count. Instead of running both models separately, total mosquito count data was selected as the singular outcome variable as it pertains better to the question on perceived mosquito abundance and its effects on human behavior. The issue of composition as it pertains to transmission risk would be better estimated through a more robust data set.

Risk factors with a univariate analysis that failed to converge or resulted in an error of numerically singular Hessian were omitted from consideration within the

GLMMNB candidate model. Tests comparing the negative binomial model to a Poisson model showed overdispersion, suggesting that negative binomial was the proper choice. Goodness of fit tests and other tests used to check the assumptions of the negative binomial model showed a lack of normality in the variances and a poor fit at the tails of the data range. The treatment of missing data included exclusion, while outliers were included because most non zeros values were indicated as potential outliers. Generalized linear models with generalized estimating equation (GEE) grouping by household were designed to assess post hoc relationships that were of interest during the data exploration phase. These models were used to decrease the chance of multicollinearity in the final multivariate model and suggest possible alternative associations that may be useful support in further work.

The multivariate analysis is based upon a GLMMNB model using the glmmTMB R package. The negative binomial distribution utilizes a variance that increases quadratically with the mean such that the model can be represented by:

$$E(Y_{ij}) = \mu_{ij}$$

$$\mu_{ij} = \beta_{0i} + \beta_1 x_{1ij} + \ldots + \beta_p x_{pij}$$

$$v = \mu + \mu^2/\varphi$$

where the mean μ is a function of the covariates, x_{ij} is a vector of the measured covariates for the *j*th count of the *i*th site, β is a vector of parameters, v is the variance structure and φ is the dispersion parameter or exponentiation of the linear predictor from the dispersion model.[67] The household identifying number is used as a random effect to adjust for correlation in repeated individuals. Time is included as month of collection indexed 1 to 12, 1 indicating months since start of sampling (1=October, 12=September). Month of

collection defined by month and year was included in the dispersion formula, which is a one-sided formula that allows dispersion to vary with fixed effects; used to account for heteroskedasticity or increased variability by month. Q-Q plots, Residual vs predicted plots, One-sample Kolmogorov-Smirnov test, dispersion tests, zero inflation tests, and AIC were used to select and evaluate these parameters for the model. Time modelled with sine and cosine curves, month 1-24, and season were all considered as alternatives to the random effect and dispersion formulas as currently specified. Inclusion of variables into the multivariate candidate models were defined by grouping variables using a conceptual framework that highlights the theorized causal relationship of mosquito count by chemosensory ques such as animal or human factors, deterrents, seasonality, climate, and environmental suitability. Within each of these categories of consideration the available parameters were considered for collinear relationships using tests for correlation between each pair of two variables. Among variables grouped within the same type, candidate models were formulated and compared using ANOVA tests in a step wise method to determine if the addition of each variable improves the fit of the model. Candidate models were also compared using AIC statistics as well as assessments of internal validity and general fit to the data. Spatial dependency across time in the residuals from the final candidate model as well as in an intercept only model was considered using semivariogram calculation and plotting of the visual representation by month.

GLMMNB as defined in accordance with the multivariate mosquito count model, was attempted to describe the relative frequency of grouped mosquitoes, by running the model with the outcome variable being species grouping; ‘primary’, ‘secondary’,

'tertiary', and 'non' vector species of importance. The primary vector species classification is defined by the number of *An. arabiensis*. The secondary classification includes *An. rufipes, An. pretoriensis, An. rivulorum, An. squamosus,* and *An. pharoensis*. Tertiary includes *An. leesoni, An. quadriannulatus, An. coustani, An. longipalpis*, and molecularly *Anopheles* genus but unidentified species mosquitoes. The 'non' category describes the remainder of the total mosquito count from that collection. These categories were determined based on literature evaluations of malaria species of importance in Zambia.[8,68] This analytical approach failed to converge as specified due to a lack of power to detect differences. These classifications were used to assess the potential for differential spatial clustering relative to abundance of mosquitoes collected. Each class was divided into above or below the 50th percentile of non-zero mosquito count abundance at a given household across time. Coordinates from the households with species counts within each class; primary, secondary, and tertiary at a given abundance level were compared to the non-vector species belonging households at that same level of abundance, using a difference in K-function approach. Similarly, a cross K-function approach was used with this division scheme to analyze the interaction between the spatial distribution of environmental features above or below the 50th percentile in representativeness in the buffer zone around each household as they relate to the primary vector species, *An. arabiensis*, at high and low abundances.

Data was electronically collected onto standardized survey forms by the MRT field team and inputted into Redcap where IRB approved personnel from the ICEMR and MRT teams can gain access. Data cleaning was done collaboratively between MRT and JHBSPH. Analysis were conducted in R version 4.1.1 using the DHARMa and

glmmTMB packages for model building and goodness of fit tests. Spatial analysis was conducted using ArcGIS version 2.8.3 and R packages geoR, gstat, and sp.[69]

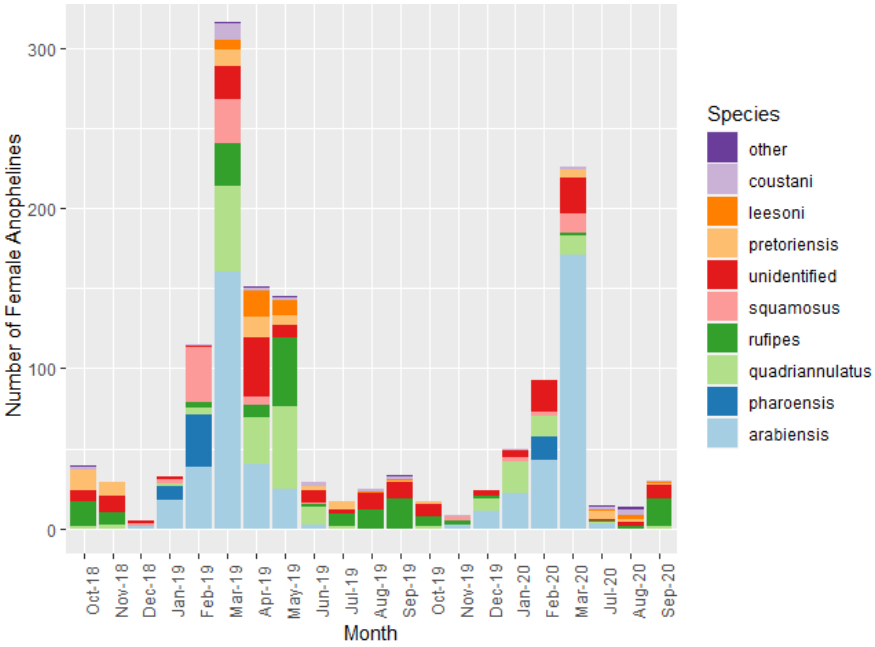
Results

Of the 202 households enrolled in the study, 59 participated in entomology collections (Figure 1). 1,116 collections were performed. A collection is defined as the field team visiting a household and setting 1 to 3 traps in a given month. Each household contributed an average of 19 months of collections over the two-year period with the maximum number of collections being 21 months, and the minimum being one month. Over the two full years of collection, 325 traps were placed in 2018 from October to December, 1,380 traps in 2019, and 650 traps in 2020. During 2020 there was a 3 month pause from April to June in which no collections were performed due to restrictions brought about by the COVID-19 pandemic. 1,113 of the 2,355 traps were placed indoors, 1,021 traps outdoors near people, and 221 outdoors near animals (See Table 1).

The number of mosquitoes collected across the study is 3,095 individuals. The mosquitoes collected and identified as culicines represent 52.5% of the overall mosquitoes collected, with a higher proportion of indoor collections being culicines 60.6%, See Appendix Table 1). Ten species of anophelines were identified through PCR analysis from the collections, the largest proportion of which were *An. arabiensis*, making up 75% of Anophelines collected indoors and 23% of those collected outdoors. The second most abundant was *An. quadriannulatus* followed *by An. rufipes*. Figure 2 represents the relative abundances of mosquitoes by species per month across the study period. Peaks in both anophelines and culicines collected occur in the months between

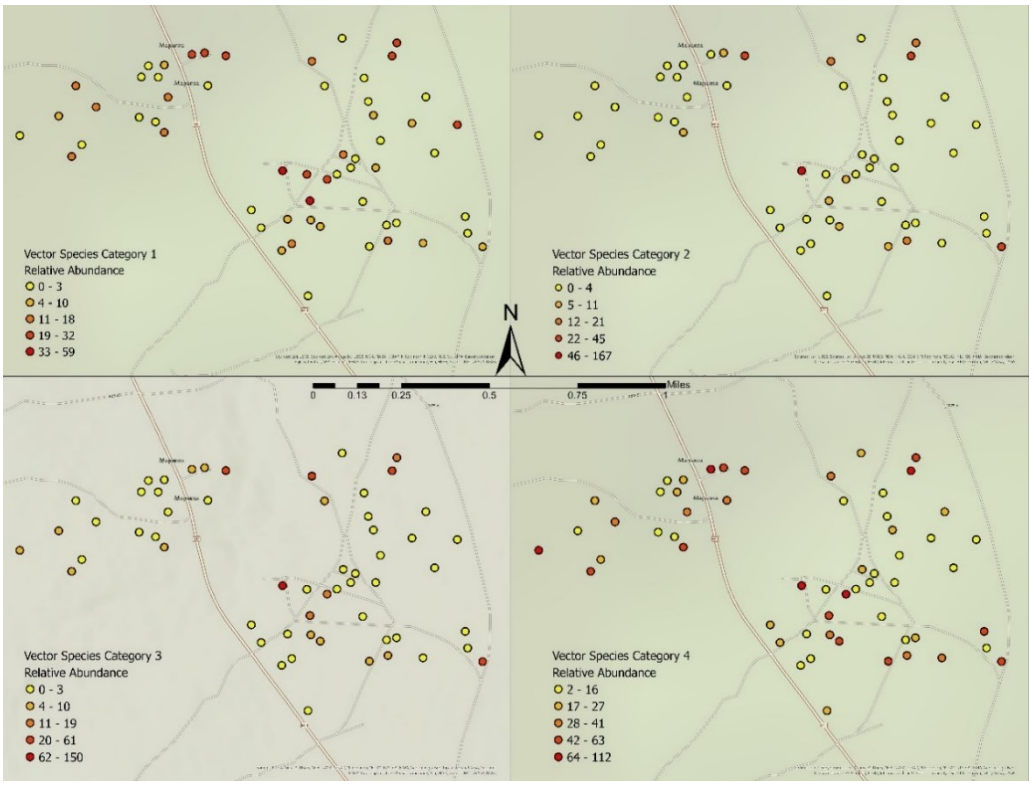
January and April of each year, falling within the rainy season (see Appendix Figure 1).

Figure 2. Number of Female Anophelines by species per month.



The relative amount of culicines to anophelines increases in the months leading to the rainy season (July to October). The relative amount of *An. arabiensis* compared to other species is higher in the rainy season as well but no significant differences exist across the relative frequencies by vector species grouping (see Appendix Figure 2). Blood meal analysis results are limited as only 50 visually blooded anopheline females were identified. Blood meal host detection was performed on all 50 visually blooded mosquitoes. Human-fed mosquitoes included nineteen *An. arabiensis* and one *An. squamosus.* One mosquito tested positive for a *P. falciparum* CSP. The positive mosquito was an *An. squamosus* collected outdoors. All of the visually blooded mosquitoes collected indoors except one *An. coustani*, were *An. arabiensis* (see Appendix Figure 3). Due to the limited number of blooded mosquitoes and CSP positive mosquitoes, no further analysis were conducted on this outcome.

Figure 3. Mosquito abundance across Mapanza, Choma District, Zambia separated by competence as a malaria vector.



The potential for spatial clustering amongst different categories of malaria-vector species' was considered using a difference in K-function approach in which each category is compared with category 4 at preset levels of abundance being greater or less than the 50th percentile count value. The difference in the amount of clustering suggested was not significant in either range of abundance when comparing potential vector species to non-vector species across any classification of vector competence or abundance (Appendix Figure 4). There is some small increase suggested in primary vector species clustering relative to non-vector species in the low abundance range particularly at the greater distances and some increase in clustering amongst non-vector species relative to secondary and tertiary vector species at the lower abundance ranges. In the upper abundance range there is low magnitude evidence of greater clustering in anopheline

species at short distances and non-anopheline species at greater distances. A cross K-function approach examining the potential for spatial interaction between environmental features and the incidence of *An. arabiensis* collection, divided by greater or less than the 50th percentile levels of abundance showed that across all comparisons, the observed cross K-function is distinct from what would be expected from independence in the observations (Appendix Figure 5). This is expected given collections occur at preset locations. There does not appear to be a clear trend by abundance or land use percentile.

Household characteristics collected across the study are summarized by trap location and seasonality in Table 1. Pearson chi squared values were calculated to compare across factor levels within trap location parameters and across seasons. Significance is based on an alpha of 0.05 and is indicated with an asterisk in the top right-hand corner of the cell. Several variables were not expected to change across season due to being defined as time fixed measures, changes then would represent missingness in the data. Household level variables for example, are not expected to be any different across trap locations just as time fixed data do not change across seasons. Numbers and percentages for categorical variables and number, means, and standard deviations for continuous variables are included to show variable response across the study. Variables only collected at one location have NA's within the cells of the inapplicable trap location columns. The household contribution column depicts how many unique households of the 59 considered, contributed data to that variable field. Notably, only 3 households reported not using nets the night prior to answering the survey on initial visits and all households report owning nets. The time spent outside at night was considered as a possible predictor of net use. Using a GEE linear regression of net use on time of last entry at night for a

possible association, no significant relationship was observable (B1=0.07, p-value=0.19). Correlation matrices suggest statistically significant correlations between all measures of net use and ownership, and as such only one measure can be included in the multivariate analysis.

Changes in seasonal distribution of risk factors prompted the use of simple linear regressions to examine the difference in aggregated mosquito counts by season. The coefficients demonstrate an increase or decrease in mosquito count between the rainy and dry season across a one unit increase of the risk factor. Variables with a significant relationship (alpha = 0.05) from this set of univariate analysis include animal, intervention, and household variables. An increase in the number of livestock near a trap and the presence of other animals within five meters of the trap are both associated with an increasingly greater mosquito count in the rainy season relative to the dry season (See Appendix Table 2). This relationship holds for the number of goats owned by the household above a value of 10 goats, home ownership of 0 to ten goats has a coefficient value near the null of no difference (B1 = 0.02, B2 = 0.06). The protective impact of a fire burning near the trap is greater in the dry season as the increase in rainy season mosquitoes counts versus dry season counts is negative when a fire is set versus not (B1=12.66, $p<0.05$). A similar relationship exists for living within 50 meters of resting water, suggesting that dry season counts are more substantially affected by natural water sources. Several household variables such as roofing and door types as well as the presence of holes in the structure a trap was placed within, are significantly associated with a change in relative magnitude by season, dependent upon the material class (See Appendix Table 2).

Table 1. Baseline Characteristics of Household Collections by Trap Location

	Trap Location			Collection Seasonality	
	Indoor Collection (n=1113)	Outdoor Trap by Humans (n=1021)	Outdoor Trap by Animals (n=221)	Dry Season	Rainy Season
Number of Goats:	*	*	*		
0-#(%)	796(74.5)	723(73.8)	27(13.3)	699(68.7)	847(68.6)
1-5-#(%)	68 (6.36)	63(6.43)	20(9.85)	55(5.4)	96(7.8)
5-10-#(%)	93 (8.70)	90(9.18)	69(33.99)	115(11.3)	137(11.1)
10-20-#(%)	75 (7.02)	70(7.14)	52(25.62)	93(9.14)	104(8.4)
>20-#(%)	37 (3.46)	34(3.47)	35(17.24)	55(5.41)	51(4.1)
					P=0.14
Number of Chickens:	*	*	*		*
0-#(%)	362(34.44)	362(33.83)	45(22.50)	292(28.74)	409(33.12)
1-5-#(%)	205(19.51)	205(19.16)	36(18.00)	196(19.29)	248(20.08)
5-10-#(%)	131(12.46)	140(13.08)	36(18.00)	100(9.84)	207(16.76)
10-20-#(%)	151(14.37)	157(14.67)	48(24.00)	169(16.63)	187(15.14)
>20-#(%)	202(19.22)	206(19.25)	35(17.50)	259(25.49)	184(14.90)
					P=<0.001
Fire burning near trap:	*	*	*		*
No-#(%)	895(85.6)	34(3.6)	171(94.0)	462(47.5)	638(53.6)
Yes-#(%)	151(14.4)	902(96.4)	11(6.0)	511(52.5)	553(46.4)
					P=0.005
Prior Elephantiasis treatment:					
No-#(%)	371(35.4)	409(36.2)	92(42.8)	408(36.6)	464(36.4)
Yes-#(%)	677(64.6)	720(63.8)	123(57.2)	708(63.4)	812(63.6)
					P=0.921
Time last person entered house at night: n,mean(sd)	1075, 20.6(2.97)	982, 20.6(2.86)	201, 20.8(2.70)	1021, 20.6(3.01)	1237, 20.7(2.81)
Time of first exit in the morning: n,mean hour(sd)	1075, 5.7(0.677)	982, 5.65(0.66)	201, 5.47(0.66)	1018, 5.74(0.68)	1240, 5.56(0.65)
Duration of time spent outdoors at night: n,mean hour(sd)	1071, 2.88(1.46)	978, 2.90(1.40)	201, 3.05(1.23)	1015, 2.88(1.46)	1235, 2.92(1.37)
Time of outdoor gathering: n,mean hour(sd)	1073, 18.2(1.10)	983, 18.2(1.03)	203, 18.2(0.95)	1017, 18.2(1.14)	1242, 18.2(0.98)
Resting Water 50m from the home:					
No-#(%)	823(76.5)	756(76.7)	157(77.3)	769(75.3)	967(77.7)
	253(23.5)	230(23.3)	46(22.7)	252(24.7)	277(22.3)

Yes-#(%)					P=0.177
Source of drinking water:					
Piped-#(%)	213(18.87)	182(17.37)	44(20.47)	206(18.4)	233(18.26)
Borehole-#(%)	847(75.02)	806(76.91)	152(70.70)	844(75.62)	961(75.31)
Open Well-#(%)	69(6.11)	60(5.73)	19(8.84)	66(5.91)	82(6.43)
					P=0.870
Wall Material:	*	*	*		
Natural-#(%)	42(3.8)	41(4.0)	1(0.5)	40(3.6)	44(3.5)
Brick-#(%)	753(67.9)	701(68.2)	159(74.0)	756(69.0)	858(68.3)
Concrete-#(%)	314(28.3)	286(27.8)	55(25.6)	300(27.4)	355(28.2)
					P=0.887
Roof Material:					
Grass/Thatch-#(%)	121(10.9)	110(10.7)	24(11.6)	110(10.9)	137(10.9)
IronSheet-#(%)	943(85.0)	879(85.4)	187(87.0)	937(85.5)	1072(85.3)
Asbestos-#(%)	45(4.1)	40(3.9)	3(1.4)	40(3.6)	48(3.8)
					P=0.976
Floor Type:					
Natural-#(%)	477(43.0)	435(42.3)	83(38.6)	468(42.7)	527(41.9)
Finished-#(%)	632(57.0)	594(57.7)	132(61.4)	628(57.3)	730(58.1)
					P=0.704
Door Type:	*	*	*		
SolidWood-#(%)	138(13.0)	126(12.8)	42(20.0)	147(13.9)	159(13.2)
IronSheet-#(%)	258(24.2)	250(25.3)	107(50.7)	282(26.8)	333(27.5)
WoodPlank-#(%)	669(62.8)	611(61.9)	62(29.4)	625(59.3)	717(59.3)
					P=0.822
Eaves are:					
Closed-#(%)	536(48.3)	498(48.4)	99(46.0)	527(48.1)	606(48.2)
Open-#(%)	462(41.7)	427(41.5)	104(48.4)	464(42.3)	529(42.1)
Patially Open-#(%)	111(10.0)	104(10.1)	12(5.6)	105(9.6)	122(9.7)
					P=0.989
Structure has holes:					
No-#(%)	229(20.6)	210(20.4)	27(12.6)	209(19.1)	257(20.4)
Yes-#(%)	880(79.4)	819(79.6)	188(87.4)	887(80.9)	1000(79.6)
					P=0.403
Holes blocked at night:	*		*		
No-#(%)	469(42.3)	434(42.2)	102(47.4)	466(42.5)	539(42.9)
Yes,all-#(%)	187(16.9)	184(17.9)	56(26.0)	198(18.1)	229(18.2)
Yes,some-#(%)	453(40.8)	411(39.9)	57(26.5)	432(39.4)	489(38.9)
					P=0.968

Wind speed in miles/ second: n,mean(sd)	1113, 1.10(0.70)	1021, 1.10(0.70)	221, 1.10(0.70)	2006, 1.23(0.76)	1112, 0.94(0.56)
Rain in mm: n,mean(sd)	1113, 0.09(0.14)	1021, 0.09(0.14)	221, 0.12(0.16)	1667, 7.2e-5 (2.7e-4)	1734, 0.177 (0.164)
Temperature in Celsius: n,mean(sd)	1113, 22.5(4.62)	1021, 22.4(4.63)	221, 22.6(4.37)	2232, 21.3(5.51)	2340, 23.6(3.16)
Relative humidity in pcr: N,mean(sd)	1113, 54.8(19.4)	1021, 54.4(19.2)	221, 58.0(19.5)	2006, 41.6(10.8)	1644, 71.2(14.5)
Livestock 5m from trap: No- #(%) Yes- #(%)	NA	* 634 (68.7) 288 (31.2)	* 6 (3.24) 179 (96.8)	289(57.6) 213(42.4)	351(58.0) 254(50.0) P=0.880
Location of trap near people: Kitchen- #(%) Porch-#(%) Under Tree- #(%) Other-#(%)	NA	* 778(82.1) 117(12.3) 39(4.1) 14(1.5)	NA	354(80.4) 57(13.0) 16(3.6) 13(3.0)	424(83.5) 60(11.8) 23(4.5) 1(0.2) P=0.004
Openings present: No-#(%) Yes-#(%)	350(33.9) 682(66.1)	NA	NA	158(33.8) 309(66.2)	192(34.0) 373(66.0) P=0.960
Number sleeping in house: n,mean(sd)	1047, 3.97(2.36)	NA	NA	473, 3.9 (2.3)	574, 4.03(2.41)
Number of netshanging :n,mean(sd)	1048, 1.04(5.46)	NA	NA	473, 0.98(5.74)	575, 1.09(5.22)
Number of people using ITN last night: n,mean(sd)	521, 2.47(1.39)	NA	NA	208, 2.27(1.28)	213, 2.60(1.45)
House sprayed within 6 months No-#(%) Yes-#(%)	832(81.3) 191(18.7)	NA	NA	410(87.6) 58(12.4)	* 422(76.0) 133(24.0) P=<0.001

#=number, %=percent, sd=standard deviation, n=sample size, p= p-value

The univariate analysis using GLMMNB indicates change in risk of higher or lower mosquito counts relative to change in the unit of the risk factor variable. This

relationship is evaluated on the monthly visit time scale (See Appendix Table 3). Variables showing significant relationships with mosquito count (alpha of 0.1) include the number of goats owned by the household with a greater strength of association in the higher goat count categories, the number of people sleeping in the structure where the trap was placed, the number of ITN nets hanging in the house where the collection was made, the number of people using an individual ITN in the household, whether or not the house had been sprayed with insecticide within the last six months, the time individuals from the household started to gather outside at night, the location type of the trap, the primary water source for the household, if eaves are open, the average rainfall in mm, temperature in Celsius, and humidity in pct, and the percent of landcover around the household made up of trees. Windspeed is one of few variables with a protective effect (RR=0.093, p-value<0.005). Time of outdoor gathering indicates later gathering times, which are shown to correlate with shorter outdoor gathering durations, are associated with a higher relative risk of greater mosquito abundance.

Correlation matrices between variables of a given type determined that all variables pertaining to animals and that indicate a significant association with mosquito count are statistically significantly correlated (alpha = 0.05). Amongst intervention and behavioral variables net use and fire burning are significantly correlated. All variables describing household structure are significantly correlated. All measures of climate are significantly correlated. The results from Table 1 and Table 2 prompted the analysis of the effect of climate on behavioral and intervention variables using generalized linear models. The only intervention that indicated a slight although significant relationship with the climate variables is time spent outside at night with humidity (B=0.01, P-

value=0.048). The inclusion of time of day as a modifier to these relationships did not have an effect on the estimates.

Fixed effect estimates from the multivariate analysis constructed upon inclusion of variables from each category of risk factor, selected based upon the strengths of association with mosquito count and their optimization in data availability is presented in Table 2 as 'Multivariate Full', with a full presentation of results in Appendix Table 4. A subsetted version of this model with only one predictor from each category, selection based on significance, is also included. A likelihood ratio test suggests that the extended set of variables do not meaningfully contribute to the model (chi sq p-value = 0.79). The subsetted model suggests goat ownership is associated with increased risk of high mosquito abundance and that this relationship is greater, the more goats owned by the household. Trees coverage around the home is associated with a slight but significant increase, later times of outdoor gathering is protective, and open well and bore hole drinking water sources are more positively associated with increased risk compared to piped water. The number of ITN nets hanging in the home is also significantly associated with a higher relative risk.

Table 2. GLMMNB Main Effect Model

	Univariate		Multivariate Full		Multivariate Subset	
	RR	p-value	RR	p-value	RR	p-value
Number of Goats						
0 Goats=ref	4.738	0.000				
1-5 Goats	1.582	0.077	1.500	0.140	1.447	0.170
5-10 Goats	2.990	0.000	2.709	0.001	2.426	0.001

10-20 Goats	2.670	0.001	2.881	0.000	2.694	0.000
>20 Goats	6.888	0.000	8.903	0.000	9.392	0.000
Percent of landcover around household made up of trees						
(Intercept)	4.536	0.000				
% Tree	1.067	0.065	1.076	0.015	1.046	0.027
Time of outdoor gathering						
(Intercept)	83.19	0.000				
Time Outdoors	0.859	0.015	0.890	0.091	0.882	0.037
Source of drinking water						
Piped=ref	4.914	0.000				
Bore Hole	1.164	0.613	1.299	0.283	1.258	0.325
Open Well	4.321	0.009	2.986	0.025	2.968	0.025
Number of nets hanging						
(Intercept)	4.268	0.000				
# of ITN	1.253	0.003	1.282	0.002	1.270	0.002
Number sleeping in house						
(Intercept)	3.957	0.000				
# Sleeping in HH	1.066	0.078	1.020	0.563		
House sprayed within 6 months						
No = ref	4.711	0.000				
Yes	1.327	0.093	1.003	0.503		
Distance to the nearest water body						
(Intercept)	8.060	0.000				
Distance to Water	0.999	0.107	1.001	0.458		
Fixed Effect						
(Intercept)			10.88	0.446	37.12	0.004

(Month 1-12)		0.935	0.068	0.935	0.065

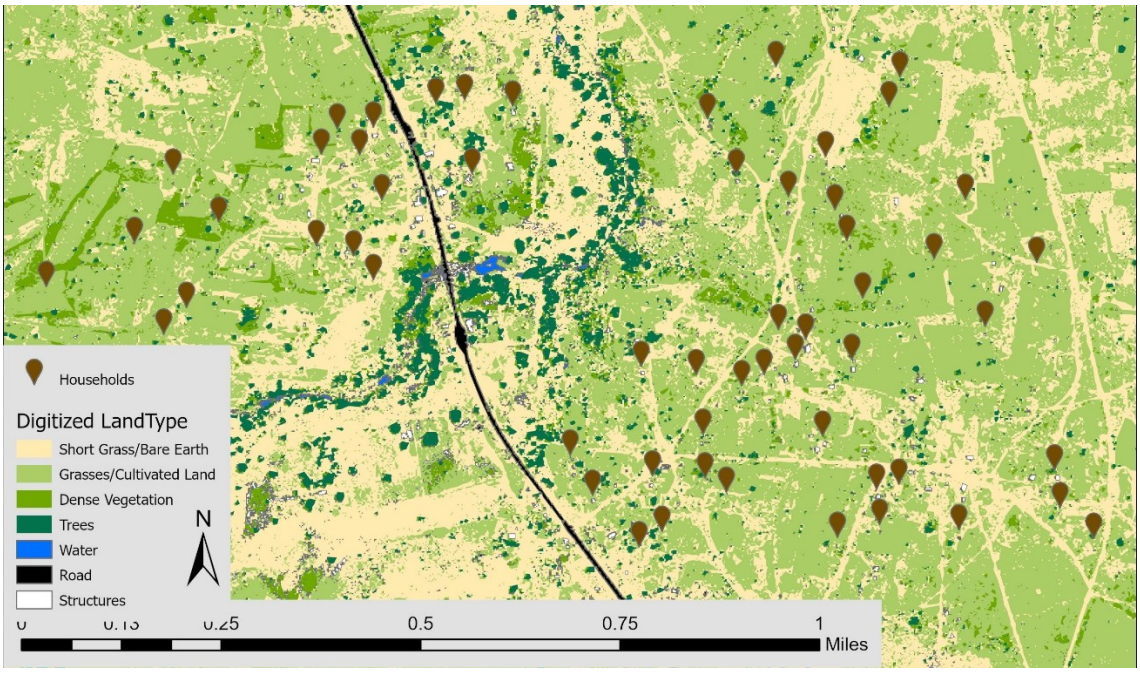
HH = household, % = percent, # = number

Across time, allowing for variable dispersion by month, greater variation of mosquito counts are seen largely in the rainy season (November to April), although this trend is most evident from January to May in the first year of study (Appendix Table 4). In the second year of study dispersion is greatest November to March, and cannot extend further due to the pause in collection from April to June. Testing for temporal autocorrelation through aggregating the residuals by household across time using the Durbin-Watson test was nonsignificant (p=0.11), with increases and decreases in projected autocorrelation corresponding to the months with the greatest variation in counts. This result suggests that the flexibility with time built into the model is properly informed. Spatial autocorrelation was tested through aggregating repeated observations by location and using the Morans I test (p=0.47) (Appendix Figure 6). Spatial dependency in the residuals from the above model is considered per individual month, across the study through the use of semivariogram figures (Appendix Figure 7). These figures suggest that there is spatiotemporal variation not accounted for in the developed model. Q-Q plots of the residuals showed a good fit to the data. The one-sample Kolmogorov-Smirnov test (p=0.14), dispersion test (p=0.83), and Outlier test (p=0.58) were all nonsignificant. The residuals versus predicted plot divided into quantiles showed good fit (Appendix Figure 8).

Of the environmental variables constructed using a supervised land use classification scheme (Figure 4), only percent tree cover was significantly associated with mosquito count in the univariate analysis. Although insignificant, a multivariate model

was designed to explore the association between mosquito abundance and climatic variable with adjustment (Appendix Table 5). Higher risk with higher temperature is the only significant risk factor; higher relative abundance of trees and dense vegetation are also associated but not statistically significant (RR = 1.0, p=0.25 and RR=1.04, p=0.11 respectively). A multivariate model to describe intervention use with adjustment is also considered. Time of outdoor gathering and number of nets hanging in the household are the only significant risk factors. Later time of outdoor gathering and time of first exit in the morning are the only risk factors associated with a decreased relative risk. The other interventions and behaviors are null in effect, except number of nets hanging which is associated with increased risk.

Figure 4. Land use map of Mapanza, Choma District, Zambia featuring contributing households.



Discussion

Overview

Low malaria transmission settings continue to sustain the spread of disease at low levels, have the potential for surges in cases and circulating parasites given the existence of viable hosts and environmental features to support arthropod-host lifecycles, and are potentially less typical in response to interventions designed for impact in high transmission settings. Several challenges exist, however, in analyzing entomological data from larger epidemiological studies in low malaria transmission settings due to inherent limitations relating to study sampling design, quality and completeness of data, small numbers of observations, and low specimen counts. Household level factors associated with mosquito abundance were examined using a multilevel modeling framework that produced information on strength of association, significance, and direction for five fields of risk factor information: animal ownership, environment features, human behavior, household features, and use of interventions. Vector species abundance and composition showed no significant differences across space or relative abundance when analyzed using cross k-function and differences in k-function approaches. This is not surprising due to the fine geographic scale. Vector species abundance analysis was not powered to analyze associations with household level risk factors using a multilevel approach, entomologically relevant inferences are able to be made however through the exploration of the species data. Associations between mosquito abundance and human behaviors relevant to transmission such as time spent outdoors and intervention use was analyzed using a multilevel approach, although the only significant risk factors were also included in the main model with little to no change in the association presented. Similarly, a model

was developed to analyze environmental features' effect on abundance, controlling for the effect of other features which resulted in null effects by all factors except temperature, likely due to strong correlations existing between these predictors.

Entomological Inferences

Within in this study, ten different species of anophelines were identified, the most abundant of which was *An. arabiensis* which is consistent with prior data from this area.[82268] *Anopheles arabiensis* demonstrates a high correlation with the rainy season, provided a one-month lag as counts begin to increase in December and wane by May. This pattern is also visible in *An. pharoensis* and *An. squamosus* while *An. rufipes* shows relatively consistent abundances across the year, with a slight rise in the dry season. This data suggests that seasonality is an important aspect in overall abundance as well as the composition of species at any given time. Notably, there is a complete absence of *An. funestus*, a species reported in Choma District, Zambia prior to extended droughts in 2005.[68] Species composition in terms of vector competence is another important aspect to consider, although this requires a greater number of specimens to be collected in order to find discernable differences by space, time, or association with known risk factors. Spatial analysis considering abundance of different classes of vector species by competence did not reveal any statistically significant results with respect to tendency to cluster spatially or tendency to exist as modifiable by the presence or absence of environmental features.

In addition to the ten species identified, 106 specimens were morphologically or molecularly unidentifiable due to the poor condition of the specimen. Specimens may

also have been unidentifiable due to limitations in the ITS2 PCR approach and associated key for known band lengths of anopheline species. This is further supported by the 83 specimens classified as "500 bp" in Appendix Table 1 that were not able to be classified to species as this band size is associated with multiple species. Of the mosquitoes collected, only one was CSP-positive which demonstrates the low infectivity of mosquitoes in this low transmission area, but may also be influenced by the relatively small number of mosquitoes analyzed in this study. Analysis for infection was only performed on visually blood-fed mosquitoes of which there were only 50. Restricting analysis to visually blooded mosquitoes has the potential for missing specimen that have partially digested or have small blood meals in the midgut and may be CSP-positive.[70]

Model Inferences

Univariate and multivariate approaches to describe the association of mosquito counts with known risk factors using negative binomial generalized linear mixed-effect modeling showed significant association with goat ownership, tree coverage around the home, time of outdoor gathering, source of drinking water and number of nets hanging in the home. Many risk factors expected to be strongly associated with mosquito were insignificant, such as several household structural elements and interventions. All interventions examined with the exception of the number of nets hanging were nearly null while net use was actually associated with a slight increase in expected risk of higher counts. It is indiscernible if species composition is a major factor in attenuating the observed relationships as susceptibility to the insecticidal and repellency effects of net and IRS use are not evenly felt by different mosquito species.[3,24] It is also possible that

greater rates of intervention usage is due to higher perceived and experienced mosquito burden which would then attenuate the protective effect of the intervention relative to other times or places where both mosquito burden and as a results, net usage is lower. To determine this, data on mosquito abundance before intervention access would need to be assessed to understand the true counterfactual. Given a strong effect on mosquito burden, as observed in higher transmission settings these considerations have not been necessary for associations to be statistically significant.[7,71] Accordingly this result suggests a need for more effective interventions to a variety of mosquito species including cryptic vector species in low transmission settings.

Other results from the GLMMNB main model that should prompt further analysis include later times of outdoor gathering being associated with higher mosquito abundance risk, this is of concern since peak biting times of *An. arabiensis* start in the later part of the night at around 10:00 pm.[68] In terms of drinking water sources, bore hole and open well sources both show an increased risk of high mosquito burden as compared to piped water which is likely due to the relatively open nature of water sources around the home providing oviposition sites. Boreholes have been implemented as improvements to hygiene by WASH, a UN effort, but have also been linked to other health events relating to water safety.[72] Higher relative tree coverage being associated with higher mosquito counts is likely also due to providing favorable conditions for harborage of the mosquitoes either before or after they feed. While microclimate data was not collected, weather data obtained from a weather station near the catchment area was used to make inferences on household conditions across time in conjunction with the evaluation of environmental features and household structure data. This analysis suggested that higher

temperatures are associated with greater risks of higher mosquito abundances, aligning with prior work based upon the conditions affecting the rate of development of mosquitoes.[37,73]

Incorporation of time in a dispersion formula allowed for the demonstration of shifts in variability of mosquito counts by month which followed a seasonal trend. The rise in dispersion followed the increase in the upper levels of mosquito counts or the increases in populations existing within favorable conditions starting in December and extending to May. The rainy season in Choma, Zambia is November to April, which supports the increase of mosquito by season following a one-month lag. The one-month lag is the approximate time it takes for a generation of mosquitoes to develop from eggs to adults capable of foraging and being caught in a trap.[74]

Limitations

Probability sampling by geographic area rather than incorporating active case detection and focal transmission used to recruit participants presents a limitation; inability to link cases with entomological data and lower sensitivity within a low transmission setting resulting in a likely inflated lack of positives which does not allow for analysis of transmission, rather the potential for transmission given an introduction of parasites into the areas surrounding the households under study. In addition, limitations exist in the ability to study particular measures due to non-response in survey data. This was handled by optimizing the analytical data set to exclude missing observations but maintain the greatest number of observations and retain similar mean and variance estimates relative to the full dataset. Multiple imputation methods were considered, but

the potential for this approach to produce biased estimates towards the null was more likely relative to excluding observations with missing data after optimizing the analytical dataset revealed less than 20% of observations were excluded and the mean and variance estimates changed by less than 10%.

Low counts of species resulted in issues with singularity and convergence of models and as such results were not reportable. Additional approaches to sampling should be taken in future work to develop analytical datasets with sufficient power to examine these risk factor relationships. This study only used non baited CDC light traps which has the potential to restrict the majority of collections to foraging mosquitoes near the home, rather than fed mosquitoes. Using human landing catch techniques as well as aspirating could be useful in providing a better estimate of burden as well as more heavily powered blood meal and infection analysis results.

Future Work

For future work, space-time Bayesian approaches may be better suited to a study design of this type, since there are spatiotemporal factors that are not accounted for, as evidence by the monthly semivariograms. Since the goals of this study is inferential in nature, this is not a concern to validity but would be more necessary if extending the objectives to prediction. Kriging is also an approach that could be used in the event prediction is a goal. Semivariogram construction by time was described in order to understand the changes in dependence of observations by space across time, although in the event predictive models on this type of data are supported, this approach would be

used in developing a universal kriging approach that could be used to account for the dependence structure and provide for better predictive accuracy.

Conclusion

Despite the limitations in the data, the main GLMMNB model performed well under diagnostic examination and produced associations that are substantiated by prior work but also indicated the need for improved integrated vector control to reduce the risks of transmission as the area approaches elimination. Household and environmental feature data such as animal ownership, water sources creating breeding sites, and tree coverage can be used in developing more personalized, context specific approaches. The lack of protective effect by interventions, however, is of major concern since targeted approaches only work if the intervention is effective.[3,22,24]

References

1. Fornadel C, Norris L, Glass G, et al. Analysis of *Anopheles arabiensis* blood feeding behavior in southern Zambia during the two years after introduction of insecticide-treated bed nets. The American Journal of Tropical Medicine and Hygiene. 2010;83(4):848-853. https://www.ncbi.nlm.nih.gov/pubmed/20889878. doi: 10.4269/ajtmh.2010.10-0242.

2. Zhong D, Hemming-Schroeder E, Wang X, et al. Extensive new *Anopheles* cryptic species involved in human malaria transmission in western Kenya. Scientific Reports. 2020;10(1):16139. https://www.ncbi.nlm.nih.gov/pubmed/32999365. doi: 10.1038/s41598-020-73073-5.

3. WHO. World Malaria Report. 2020. https://www.who.int/publications/i/item/9789240015791. Accessed 11 Jan, 2021.

4. Weiss D, Lucas T, Nguyen M, et al. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum,* 2000–17: a spatial and temporal modelling study. The Lancet (British edition). 2019;394(10195):322-331. https://dx.doi.org/10.1016/S0140-6736(19)31097-9. doi: 10.1016/S0140-6736(19)31097-9.

5. Bhatt S, Weiss D, Cameron E, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. Nature (London). 2015;526(7572):207-211. https://www.ncbi.nlm.nih.gov/pubmed/26375008. doi: 10.1038/nature15535.

6. Monitoring and Evaluation of the Global Technical Strategy for Malaria 2016–2030 and Action and Investment to defeat malaria 2016–2030. WHO. September 2016.

7. Pinchoff J, Hamapumbu H, Kobayashi T, et al. Factors associated with sustained use of long-lasting insecticide-treated nets following a reduction in malaria transmission in southern Zambia. The American Journal of Tropical Medicine and Hygiene. 2015;93(5):954-960. https://www.ncbi.nlm.nih.gov/pubmed/26324729. doi: 10.4269/ajtmh.15-0093.

8. Deutsch-Feldman M, Hamapumbu H, Lubinda J, et al. Efficiency of a malaria reactive test-and-treat program in southern Zambia: A prospective, observational study. The American Journal of Tropical Medicine and Hygiene. 2018;98(5):1382. doi: 10.4269/ajtmh.17-0865.

9. Loewenberg S. Zambia's drive to eliminate malaria faces challenges. Bulletin of the World Health Organization. 2018;96(5):302-303. https://www.ncbi.nlm.nih.gov/pubmed/29875513. doi: 10.2471/BLT.18.020518.

10. WHO. Malaria surveillance, monitoring & evaluation: a reference manual. Geneva: World Health Organization. 2018.

11. Southern & Central Africa ICEMR. Malaria transmission and the impact of control efforts in southern and central Africa. NIAID NIH Web site.

https://www.niaid.nih.gov/research/southern-central-africa-international-center-excellence-malaria-research. Updated 2020.

12. Ashton R, Prosnitz D, Andrada A, et al. Evaluating malaria programmes in moderate- and low-transmission settings: practical ways to generate robust evidence. Malaria Journal. 2020;19(1):75. https://www.ncbi.nlm.nih.gov/pubmed/32070357. doi: 10.1186/s12936-020-03158-z.

13. Rajakaruna R, Alifrangis M, Amerasinghe P, et al. Pre-elimination stage of malaria in Sri Lanka: assessing the level of hidden parasites in the population. Malaria Journal. 2010;9(1):25. https://www.ncbi.nlm.nih.gov/pubmed/20089157. doi: 10.1186/1475-2875-9-25.

14. WHO. A framework for malaria elimination. Geneva: World Health Organization. 2017.

15. Guerra C, Kang S, Citron D, et al. Human mobility patterns and malaria importation on Bioko Island. Nature Communications. 2019;10(1):2332. https://www.ncbi.nlm.nih.gov/pubmed/31133635. doi: 10.1038/s41467-019-10339-1.

16. Moss W. Southern & Central Africa ICEMR. National Institute of Allergy and Infectious Diseases website. https://www.niaid.nih.gov/research/southern-central-africa-international-center-excellence-malaria-research. Updated 2020. Accessed 11/11/21, .

17. Moonen B, Cohen J, Snow R, et al. Operational strategies to achieve and maintain malaria elimination. The Lancet (British edition). 2010;376(9752):1592-1603. https://www.clinicalkey.es/playcontent/1-s2.0-S014067361061269X. doi: 10.1016/S0140-6736(10)61269-X.

18. Ulrich J, Naranjo D, Alimi T, et al. How much vector control is needed to achieve malaria elimination? Trends in Parasitology. 2013;29(3):104-109. https://www.clinicalkey.es/playcontent/1-s2.0-S1471492213000044. doi: 10.1016/j.pt.2013.01.002.

19. Kelly-Hope L, McKenzie F. The multiplicity of malaria transmission: a review of entomological inoculation rate measurements and methods across sub-Saharan Africa. Malaria Journal. 2009;8(1):19. https://www.ncbi.nlm.nih.gov/pubmed/19166589. doi: 10.1186/1475-2875-8-19.

20. Searle K, Katowa B, Kobayashi T, et al. Distinct parasite populations infect individuals identified through passive and active case detection in a region of declining malaria transmission in southern Zambia. Malaria Journal. 2017;16(1):154. https://www.ncbi.nlm.nih.gov/pubmed/28420399. doi: 10.1186/s12936-017-1810-3.

21. Kuddus M, Rahman A. Modelling and analysis of human–mosquito malaria transmission dynamics in Bangladesh. Mathematics and Computers in Simulation. 2022;193:123-138. https://dx.doi.org/10.1016/j.matcom.2021.09.021. doi: 10.1016/j.matcom.2021.09.021.

22. Norris L, Norris D. Efficacy of long-lasting insecticidal nets in use in Macha, Zambia, against the local *Anopheles arabiensis* population. Malaria Journal. 2011;10(1):254. https://www.ncbi.nlm.nih.gov/pubmed/21880143. doi: 10.1186/1475-2875-10-254.

23. Alonso P, Besansky N, Burkot T, et al. A research agenda for malaria eradication: vector control. 2011. https://pubmed.ncbi.nlm.nih.gov/21311587/. doi: 10.1371/journal.pmed.1000401.

24. Akogbéto M, Aïkpon R, Azondékon R, et al. Six years of experience in entomological surveillance of indoor residual spraying against malaria transmission in Benin: lessons learned, challenges and outlooks. Malaria Journal. 2015;14(1):242. https://www.ncbi.nlm.nih.gov/pubmed/26063497. doi: 10.1186/s12936-015-0757-5.

25. Talman A, Ouologuem D, Love K, et al. Uptake of *Plasmodium falciparum* gametocytes during mosquito blood meal by direct and membrane feeding. Frontiers in microbiology. 2020;11:246. https://www.ncbi.nlm.nih.gov/pubmed/32194521. doi: 10.3389/fmicb.2020.00246.

26. Meibalan E, Marti M. Biology of Malaria Transmission. Cold Spring Harbor perspectives in medicine. 2017;7(3):a025452. https://www.ncbi.nlm.nih.gov/pubmed/27836912. doi: 10.1101/cshperspect.a025452.

27. Massey N, Garrod G, Wiebe A, et al. A global bionomic database for the dominant vectors of human malaria. Scientific Data. 2016;3(1):160014. https://www.ncbi.nlm.nih.gov/pubmed/26927852. doi: 10.1038/sdata.2016.14.

28. Norris L, Norris D. Phylogeny of anopheline (Diptera: Culicidae) Species in southern Africa, based on nuclear and mitochondrial genes. Journal of Vector Ecology. 2015;40(1):16-27. http://www.bioone.org/doi/full/10.1111/jvec.12128. doi: 10.1111/jvec.12128.

29. Stevenson J, Norris D. Implicating cryptic and novel *anophelines* as malaria vectors in Africa. Insects (Basel, Switzerland). 2016;8(1):1. https://www.ncbi.nlm.nih.gov/pubmed/28025486. doi: 10.3390/insects8010001.

30. Borland E, Kading R. Modernizing the toolkit for arthropod blood meal identification. Insects (Basel, Switzerland). 2021;12(1):37. https://www.ncbi.nlm.nih.gov/pubmed/33418885. doi: 10.3390/insects12010037.

31. Kirby M, West P, Green C, et al. Risk factors for house-entry by culicine mosquitoes in a rural town and satellite villages in Gambia. Parasites & Vectors. 2008;1(1):41. https://www.ncbi.nlm.nih.gov/pubmed/18939969. doi: 10.1186/1756-3305-1-41.

32. Ogoma S, Lweitoijera D, Ngonyani H, et al. Screening mosquito house entry points as a potential method for integrated control of endophagic filariasis, arbovirus and malaria vectors. PLoS Neglected Tropical Diseases. 2010;4(8):e773. https://www.ncbi.nlm.nih.gov/pubmed/20689815. doi: 10.1371/journal.pntd.0000773.

33. Khayrandish A, Wood R. A multiple basis for insecticide resistance in a strain of *Culex quinquefasciatus* (Diptera: Culicidae) from Muheza, Tanzania, studied as resistance declined. Bulletin of Entomological Research. 1993;83(1):75-85. https://dx.doi.org/10.1017/S0007485300041808. doi: 10.1017/S0007485300041808.

34. Irish S, N'guessan R, Boko P, et al. Loss of protection with insecticide-treated nets against pyrethroid-resistant *Culex quinquefasciatus* mosquitoes once nets become holed: an experimental hut study. Parasites & Vectors. 2008;1(1):17. https://www.ncbi.nlm.nih.gov/pubmed/18564409. doi: 10.1186/1756-3305-1-17.

35. Charlwood J, Vij R, Billingsley P. Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of east Africa. The American Journal of Tropical Medicine and Hygiene. 2000;62(6):726-732. https://agris.fao.org/agris-search/search.do?recordID=NL2001003836. doi: 10.4269/ajtmh.2000.62.726.

36. Mwangangi J, Mbogo C, Orindi B, et al. Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years. Malaria Journal. 2013;12(1):13. https://www.ncbi.nlm.nih.gov/pubmed/23297732. doi: 10.1186/1475-2875-12-13.

37. Thomson M, Ukawuba I, Hershey C, et al. Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. The American Journal of Tropical Medicine and Hygiene. 2017;97(3_Suppl):32-45. https://www.ncbi.nlm.nih.gov/pubmed/28990912. doi: 10.4269/ajtmh.16-0696.

38. Mattah P, Futagbi G, Amekudzi LK, et al. Diversity in breeding sites and distribution of *Anopheles* mosquitoes in selected urban areas of southern Ghana. Parasites & Vectors. 2017;10(1):25. https://www.ncbi.nlm.nih.gov/pubmed/28086941. doi: 10.1186/s13071-016-1941-3.

39. Yamana T, Eltahir E. Incorporating the effects of humidity in a mechanistic model of *Anopheles gambiae* mosquito population dynamics in the Sahel region of Africa. Parasites & Vectors. 2013;6(1):235. https://www.ncbi.nlm.nih.gov/pubmed/23938022. doi: 10.1186/1756-3305-6-235.

40. Das S, Muleba M, Stevenson J, et al. Habitat partitioning of malaria vectors in Nchelenge District, Zambia. The American Journal of Tropical Medicine and Hygiene. 2016;94(6):1234-1244. https://www.ncbi.nlm.nih.gov/pubmed/27001755. doi: 10.4269/ajtmh.15-0735.

41. Bigoga J, Nanfack F, Awono-Ambene P, et al. Seasonal prevalence of malaria vectors and entomological inoculation rates in the rubber cultivated area of Niete, South Region of Cameroon. Parasites & Vectors. 2012;5(1):197. https://www.ncbi.nlm.nih.gov/pubmed/22963986. doi: 10.1186/1756-3305-5-197.

42. Nwankwo A, Okuonghae D. Mathematical assessment of the impact of different microclimate conditions on malaria transmission dynamics. Mathematical Biosciences

and Engineering : MBE. 2019;16(3):1414-1444. https://www.ncbi.nlm.nih.gov/pubmed/30947427. doi: 10.3934/mbe.2019069.

43. Paaijmans K, Read A, Thomas M. Understanding the link between malaria risk and climate. Proceedings of the National Academy of Sciences - PNAS. 2009;106(33):13844-13849. https://agris.fao.org/agris-search/search.do?recordID=US201301669099.

44. Agusto F, Gumel A, Parham P. Qualitative assessment of the role of temperature variations on malaria transmission dynamics. Journal of Biological Systems. 2015;23(4):1550030. http://www.worldscientific.com/doi/abs/10.1142/S0218339015500308. doi: 10.1142/S0218339015500308.

45. Ohm J, Baldini F, Barreaux P, et al. Rethinking the extrinsic incubation period of malaria parasites. Parasites & Vectors. 2018;11(1):178. https://www.ncbi.nlm.nih.gov/pubmed/29530073. doi: 10.1186/s13071-018-2761-4.

46. Thomas S, Ravishankaran S, Justin N, et al. Microclimate variables of the ambient environment deliver the actual estimates of the extrinsic incubation period of P*lasmodium vivax* and *Plasmodium falciparum*: a study from a malaria-endemic urban setting, Chennai in India. Malaria Journal. 2018;17(1):201. https://www.ncbi.nlm.nih.gov/pubmed/29769075. doi: 10.1186/s12936-018-2342-1.

47. Paaijmans K, Thomas M. The influence of mosquito resting behavior and associated microclimate for malaria risk. Malaria Journal. 2011;10(1):183. https://www.ncbi.nlm.nih.gov/pubmed/21736735. doi: 10.1186/1475-2875-10-183.

48. Smith D, Battle K, Hay S, et al, Ross Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. PLoS Pathogens. 2012;8(4):e1002588. https://www.ncbi.nlm.nih.gov/pubmed/22496640. doi: 10.1371/journal.ppat.1002588.

49. Chen Y, Shiu T, Tseng L, et al. Dynamic changes in genetic diversity, drug resistance mutations, and treatment outcomes of *falciparum* malaria from the low-transmission to the pre-elimination phase on the islands of São Tomé and Príncipe. Malaria Journal. 2021;20(1):467. https://www.ncbi.nlm.nih.gov/pubmed/34906134. doi: 10.1186/s12936-021-04007-3.

50. Gatton ML, Chitnis N, Churcher T, et al. The importance of mosquito behavioral adaptation to malaria control in Africa. Evolution. 2013;67(4):1218-1230. https://api.istex.fr/ark:/67375/WNG-M20GSX2B-9/fulltext.pdf. doi: 10.1111/evo.12063.

51. Najera J, Global partnership to Roll Back Malaria. Malaria control: achievements, problems and strategies. 1999.

52. Ishengoma D, Mandara C, Francis F, et al. Efficacy and safety of artemether-lumefantrine for the treatment of uncomplicated malaria and prevalence of Pfk13 and Pfmdr1 polymorphisms after a decade of using artemisinin-based combination therapy in

mainland Tanzania. Malaria Journal. 2019;18(1):88. https://www.ncbi.nlm.nih.gov/pubmed/30898164. doi: 10.1186/s12936-019-2730-1.

53. Hamid-Adiamoh M, Nwakanma D, Assogba BS, et al. Influence of insecticide resistance on the biting and resting preferences of malaria vectors in the Gambia. PloS One. 2021;16(6):e0241023. https://search.proquest.com/docview/2544864684. doi: 10.1371/journal.pone.0241023.

54. Mnzava A, Knox T, Temu E, et al. Implementation of the global plan for insecticide resistance management in malaria vectors: progress, challenges and the way forward. Malaria Journal. 2015;14(1):173. https://www.ncbi.nlm.nih.gov/pubmed/25899397. doi: 10.1186/s12936-015-0693-4.

55. Ghebreyesus T, Lynch M, Coll-Seck A. The global malaria action plan. Roll Back Malaria Partnership. 2008. https://www.unhcr.org/4afac5629.pdf.

56. Wanzirah H, Tusting L, Arinaitwe E, et al. Mind the gap: house structure and the risk of malaria in Uganda. PloS One. 2015;10(1):e0117396. https://www.ncbi.nlm.nih.gov/pubmed/25635688. doi: 10.1371/journal.pone.0117396.

57. Jatta E, Jawara M, Bradley J, et al. How house design affects malaria mosquito density, temperature, and relative humidity: an experimental study in rural Gambia. The Lancet. Planetary Health. 2018;2(11):e498-e508. https://www.ncbi.nlm.nih.gov/pubmed/30396441. doi: 10.1016/S2542-5196(18)30234-1.

58. Kizito J, Kayendeke M, Nabirye C, et al. Improving access to health care for malaria in Africa: a review of literature on what attracts patients. Malaria Journal. 2012;11(1):55. https://explore.openaire.eu/search/publication?articleId=dedup_wf_001::9fd4be2086328c1f277ab5c58b205bf2. doi: 10.1186/PREACCEPT-2317562776368437.

59. Monroe A, Harvey SA, Lam Y, et al. “People will say that I am proud”: a qualitative study of barriers to bed net use away from home in four Ugandan districts. Malaria Journal. 2014;13(1):82. https://www.ncbi.nlm.nih.gov/pubmed/24602371. doi: 10.1186/1475-2875-13-82.

60. Wesolowski A, Eagle N, Tatem A, et al. Quantifying the Impact of human mobility on malaria. Science (American Association for the Advancement of Science). 2012;338(6104):267-270. https://www.jstor.org/stable/41703718. doi: 10.1126/science.1223467.

61. ICEMR - Macha Research Trust. https://macharesearch.org/icemr/. Updated 2019. Accessed 11/11/21.

62. Coetzee M. Key to the females of Afrotropical *Anopheles* mosquitoes (Diptera: Culicidae). Malar Journal. 2020;19(1). doi: 10.1186/s12936-020-3144-9.

63. Kent R, Norris D. Identification of mammalian blood meals in mosquitoes by a multiplexed polymerase chain reaction targeting cytochrome B. The American Journal of

Tropical Medicine and Hygiene. 2005;73(2):336-342. https://www.ncbi.nlm.nih.gov/pubmed/16103600. doi: 10.4269/ajtmh.2005.73.336.

64. Koekemoer L, Kamau L, Hunt RH, et al. A cocktail polymerase chain reaction assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. The American Journal of Tropical Medicine and Hygiene. 2002;66(6):804-811. http://www.ajtmh.org/cgi/content/abstract/66/6/804. doi: 10.4269/ajtmh.2002.66.804.

65. Scott J, Brogdon W, Collins F. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. The American Journal of Tropical Medicine and Hygiene. 1993;49(4):520-529. http://www.ajtmh.org/cgi/content/abstract/49/4/520. doi: 10.4269/ajtmh.1993.49.520.

66. Gebhardt M, Searle K, Kobayashi T, et al. Understudied anophelines contribute to malaria transmission in a low-transmission setting in the Choma District, Southern Province, Zambia. The American Journal of Tropical Medicine and Hygiene. 2022. https://www.ncbi.nlm.nih.gov/pubmed/35344932. doi: 10.4269/ajtmh.21-0989.

67. Hardin J, Hilbe J. Generalized linear models and extensions. 2nd. ed. College Station, Texas: Stata Press; 2007. http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=015579559&sequence=000002&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA.

68. Fornadel CM, Norris LC, Franco V, et al. Unexpected anthropophily in the potential secondary malaria vectors *Anopheles coustani* s.l. and *Anopheles squamosus* in Macha, Zambia. Vector Borne and Zoonotic Diseases (Larchmont, N.Y.). 2011;11(8):1173-1179. doi: 10.1089/vbz.2010.0082.

69. R Core Team, 2018.R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

70. Das S, Henning TC, Simubali L, et al. Underestimation of foraging behaviour by standard field methods in malaria vector mosquitoes in southern Africa. Malaria Journal. 2015;14(1):12. https://www.ncbi.nlm.nih.gov/pubmed/25927429. doi: 10.1186/s12936-014-0527-9.

71. Gosoniu L, Vounatsou P, Tami A, et al. Spatial effects of mosquito bed nets on child mortality. BMC Public Health. 2008;8(1):356. https://www.narcis.nl/publication/RecordID/oai:pure.rug.nl:publications%2F6ae2e4ba-e86d-4f21-9ad8-e004d7d38401. doi: 10.1186/1471-2458-8-356.

72. Kinuthia G, Gicheru M, Ngure P, et al. Lifestyles and practices that enhance malaria and typhoid fever in Njoro District, Kenya. J Community Health. 2011;37(1):224-233. https://link.springer.com/article/10.1007/s10900-011-9440-0. doi: 10.1007/s10900-011-9440-0.

73. Shaman J, Spiegelman M, Cane M, et al. A hydrologically driven model of swamp water mosquito population dynamics. Ecological Modelling. 2006;194(4):395-404. https://dx.doi.org/10.1016/j.ecolmodel.2005.10.037. doi: 10.1016/j.ecolmodel.2005.10.037.

74. Maharaj R. Life table characteristics of *Anopheles arabiensis* (Diptera: Culicidae) under simulated seasonal conditions. Journal of Medical Entomology. 2003;40(6):737-742. http://www.bioone.org/doi/abs/10.1603/0022-2585-40.6.737. doi: 10.1603/0022-2585-40.6.737.

Appendix:

Tables:

Table 1. Summary of mosquito collections by location and taxonomy				
Location	Total	Anopheline	Female Anopheline	Culicine
Indoors	952	375	371	577
Outdoors	2143	1099	784 near people 300 near animal	1049

	Indoor	Outdoor	Blooded
500 bp	11	95	2
An. arabiensis	281	256	20
An. coustani	1	30	1
An. leesoni	5	34	0
An. longipalpis	1	3	0
An. pharoensis	1	54	0
An. pretoriensis	2	72	0
An. quadriannulatus	24	186	4
An. rivulorum	0	3	0
An. rufipes	14	165	7
An. squamosus	8	83	7*
Unidentified	22	61	9

Table 2. GLM of y=rainy mosquito count – dry mosquito count per household	Estimate	P-value
(Intercept)	1.175	<0.005
factor(out_trap_livestock_num)1	1.738	0.042
factor(out_trap_livestock_num)2	2.853	<0.005
factor(out_trap_livestock_num)3	5.938	<0.005
factor(out_trap_livestock_num)4	5.184	<0.005
(Intercept)	1.286	<0.005
out_trap_oth_5m_animals	1.588	<0.005
(Intercept)	1.5900	<0.005
factor(mc_no_goats)1	0.029	0.947
factor(mc_no_goats)2	0.063	0.845
factor(mc_no_goats)3	1.601	<0.005
factor(mc_no_goats)4	2.825	<0.005
(Intercept)	1.251	<0.005
in_trap_itn_num	0.555	<0.005
(Intercept)	1.482	<0.005
in_trap_house_irs	1.106	<0.005
(Intercept)	2.799	<0.005

trap_fire	-1.665	<0.005
(Intercept)	2.102	<0.005
mc_resting_water_yn	-0.643	<0.005
(Intercept)	0.825	<0.005
factor(sq_roof)2	1.384	<0.005
factor(sq_roof)3	-1.558	0.016
(Intercept)	3.452	<0.005
factor(sq_door)2	-0.686	0.0637
factor(sq_door)3	-2.044	<0.005
factor(sq_door)99	-3.914	<0.005
(Intercept)	2.044	<0.005
factor(sq_eave)1	0.126	0.584
factor(sq_eave)2	-1.474	<0.005
(Intercept)	1.094	<0.005
sq_wall_hole	0.989	<0.005
(Intercept)	2.131	<0.005
factor(sq_wdw)1	-1.132	<0.005
factor(sq_wdw)2	-0.009	0.972

Table 3. Univariate GLMMNB Analaysis	Coefficient	RR	p-value
Number of animals in a shelter near the trap (207 observations, 22 households)			
0 = ref	2.565	13.013	0.000
1-5 Animals	-0.371	0.689	0.387
5-10 Animals	0.249	1.283	0.520
10-20 Animals	0.151	1.163	0.684
>20 Animals	-0.031	0.968	0.952
Livestock 5m from a trap placed near people: (1002 observations, 57 households)			
No = ref	1.738	5.687	0.000
Yes	0.216	1.241	0.172
Livestock 5m from a trap placed near animals: (197 observations, 19 households)			
No = ref	2.870	17.641	0.000
Yes	-0.245	0.783	0.709
Number of Goats: (1029 observations, 58 households)			
0 Goats = ref	1.556	4.738	0.000
1-5 Goats	0.459	1.582	0.077
5-10 Goats	1.095	2.990	0.000
10-20 Goats	0.982	2.670	0.001
>20 Goats	1.9293	6.888	0.000
Number of Chickens (977 observations, 58 households)			

0 Chickens = ref	1.740	5.703	0.000
1-5 Chickens	-0.112	0.894	0.567
5-10 Chickens	0.340	1.405	0.120
10-20 Chickens	0.031	1.032	0.891
>20 Chickens	-0.097	0.908	0.675
Openings present: (1105 observations, 59 households)			
No=ref	1.725	5.613	0.000
Yes	-0.032	0.968	0.829
Number sleeping in house: (1031 observations, 59 households)			
(Intercept)	1.375	3.957	0.000
in_trap_num_sleep	0.063	1.066	0.078
Number of nets hanging: (1103 observations, 59 households)			
(Intercept)	1.451	4.268	0.000
in_trap_itn_num	0.225	1.253	0.003
Number of people using ITN last night: (509 observations, 55 households)			
(Intercept)	1.359	3.894	0.000
in_trap_num_usenet	0.149	1.161	0.042
House sprayed within 6 months (1082 observations, 59 households)			
No = ref	1.549	4.711	0.000
Yes	0.283	1.327	0.093
Fire burning near indoor trap: (1106 observations, 59 households)			
No = ref	1.694	5.439	0.000
Yes	0.141	1.151	0.451
Fire burning near outdoor trap placed near humans: (994 observations, 59 households)			
No = ref	1.680	5.367	0.000
Yes	0.144	1.154	0.677
Fire burning near outdoor trap placed near animals: (191 observations, 19 households)			
No = ref	2.741	15.497	0.000
Yes	0.145	1.156	0.758
Prior Elephantiasis treatment (1037 observations, 58 households)			
No = ref	1.956	7.072	0.000
Yes	-0.282	0.754	0.268
Time last person entered house at night (1059 observations, 59 households)			
(Intercept)	2.022	7.553	0.089

mc_time_in1	-0.018	0.983	0.751
Time of first exit in the morning (1059 observations, 59 households)			
(Intercept)	3.569	35.464	0.169
mc_time_left_house1	-0.065	0.937	0.459
Duration of time spent outdoors at night: (1028 observations, 58 households)			
(Intercept)	1.710	5.529	0.000
mc_outsidetime	0.012	1.012	0.81
Time of outdoor gathering: (1057 observations, 59 households)			
(Intercept)	4.421	83.199	0.00
mc_time_outside1	-0.152	0.859	0.015
Resting Water 50m from the home: (1030 observations, 58 households)			
No = ref	1.748	5.749	0.000
Yes	-0.083	0.920	0.647
Details of traps placed near people: (1008 observations, 58 households)			
Near people in kitchen = ref	1.755	5.782	0.000
Near people in porch	-0.071	0.931	0.794
Near people under tree	0.214	1.239	0.568
Near people, other	-1.223	0.294	0.067
Close to (2-3m) from house	-0.069	0.933	0.919
Details of traps placed near animals: (197 observations, 19 households)			
Near people in kitchen = ref	2.579	13.195	0.000
Next to animals in open kraal	-0.312	0.732	0.495
Next to animals , other	0.425	1.529	0.225
Source of drinking water: (1037 observations, 58 households)			
Piped = ref	1.592	4.914	0.000
Bore Hole	0.152	1.164	0.612
Open Well	1.463	4.321	0.009
Wall Material: (1018 observations, 56 households)			
Natural = ref	1.610	5.005	0.014
Brick	0.245	1.278	0.710
Concrete	-0.056	0.945	0.934

Roof Material: (1018 observations, 56 households)			
Grass/ Thatch=ref	1.966	7.141	0.000
Iron Sheet	-0.211	0.809	0.603
Asbestos	-0.631	0.532	0.405
Floor Type: (1018 observations, 56 households)			
Natural = ref	1.886	6.596	0.000
Finished	-0.204	0.816	0.425
Door Type: (1018 observations, 56 households)			
Solid Wood = ref	1.987	7.294	0.000
Iron Sheet	0.224	1.251	0.595
Wood Plank	-0.449	0.638	0.235
Other	-0.048	0.953	0.946
Eaves are: (1018 observations, 56 households)			
Closed = ref	1.486	4.421	0.000
Open	0.485	1.625	0.059
Patially Open	0.7052	2.024	0.106
Structure has holes: (1018 observations, 56 households)			
No = ref	1.411	4.099	0.000
Yes	0.457	1.580	0.129
Holes blocked at night: (1018 observations, 56 households)			
No = ref	1.797	6.031	0.000
Yes,all	0.094	1.098	0.796
Yes,some	-0.110	0.896	0.691
Average Rainfall per month in mm (575 observations, 56 households)			
(Intercept)	0.363	1.438	0.155
Rain	2.501	12.200	0.001
Average Temperature per month in Celsius (1059 observations, 59 households)			
(Intercept)	-2.669	0.0693424	0.002
Temp	0.189	1.208	0.000
Average relative humidity per month in pct (737 observations, 56 households)			
(Intercept)	0.123	1.131	0.689
humid	0.032	1.032	0.000
Average wind speed per month in ms (1112 observations, 58 households)			
(Intercept)	3.354	28.614	0.000
wind	-2.368	0.093	0.000

Distance to the nearest water body. (1112 observations, 58 households)			
(Intercept)	2.087	8.060	0.000
near_water	-0.001	0.999	0.107
Percent of landcover around household made up of trees (1112 observations, 58 households)			
(Intercept)	1.512	4.536	0.000
tree_perc	0.065	1.067	0.065
Percent of landcover around household made up of bare land. (1112 observations, 58 households)			
(Intercept)	1.642	5.165	0.000
barren_perc	0.003	1.003	0.709
Percent of landcover around household made up of grass/shrubbery (1112 observations, 58 households)			
(Intercept)	2.348	10.463	0.000
cult_perc	-0.010	0.990	0.248
Percent of landcover around household made up of dense vegetation (1112 observations, 58 households)			
(Intercept)	1.614	5.021	0.000
veg_perc	0.024	1.024	0.245

Table 4.	Univariate			Multivariate Full			Multivariate Subset		
	Coefficient	RR	p-value	Coefficient	RR	p-value	Coefficient	RR	p-value
Number of Goats									
0 Goats = ref	1.556	4.738	0.000						
1-5 Goats	0.459	1.582	0.077	0.406	1.500	0.140	0.370	1.447	0.170
5-10 Goats	1.095	2.990	0.000	0.997	2.709	0.001	0.886	2.426	0.001
10-20 Goats	0.982	2.670	0.001	1.058	2.881	0.000	0.991	2.694	0.000
>20 Goats	1.930	6.888	0.000	2.186	8.903	0.000	2.240	9.392	0.000
Percent of landcover around household made up of trees									
(Intercept)	1.512	4.536	0.000						
tree_perc	0.065	1.067	0.065	0.073	1.076	0.015	0.045	1.046	0.027
Time of outdoor gathering									
(Intercept)	4.421	83.199	0.000						
mc_time_outside1	-0.152	0.859	0.015	-0.116	0.890	0.091	-0.126	0.882	0.037
Source of drinking water									
Piped = ref	1.592	4.914	0.000						

Bore Hole	0.152	1.164	0.613	0.261	1.299	0.283	0.229	1.258	0.325
Open Well	1.464	4.321	0.009	1.094	2.986	0.025	1.088	2.968	0.025
Number of nets hanging									
(Intercept)	1.451	4.268	0.000						
in_trap_itn_num	0.225	1.253	0.003	0.249	1.282	0.002	0.239	1.270	0.002
Number sleeping in house									
(Intercept)	1.375	3.957	0.000						
in_trap_num_sleep	0.064	1.066	0.078	0.020	1.020	0.563			
House sprayed within 6 months									
No = ref	1.550	4.711	0.000						
Yes	0.283	1.327	0.093	0.003	1.003	0.503			
Distance to the nearest water body									
(Intercept)	2.087	8.060	0.000						
near_water	-0.001	0.999	0.107	0.001	1.001	0.458			
Fixed Effect									
(Intercept)				2.387	10.880	0.446	3.614	37.127	0.004
(Month 1-12)				-0.067	0.935	0.068	-0.067	0.935	0.065
Random Effects									
Variance				0.250			0.255		
Std dev				0.500			0.505		
Dispersion Model									
(Intercept)				-1.223	0.294	0.000	-1.211	0.298	0.000
month18-Nov				-0.432	0.649	0.303	-0.437	0.646	0.298
month18-Dec				-0.642	0.526	0.130	-0.656	0.519	0.122
month19-Jan				0.518	1.678	0.210	0.505	1.658	0.220
month19-Feb				1.142	3.134	0.008	1.120	3.065	0.009
month19-Mar				1.145	3.144	0.007	1.093	2.983	0.009
month19-Apr				1.110	3.034	0.008	1.128	3.088	0.007
month19-May				0.669	1.952	0.124	0.684	1.982	0.116
month19-Jun				-0.417	0.659	0.361	-0.429	0.651	0.346
month19-Jul				-1.547	0.213	0.004	-1.564	0.209	0.003
month19-Aug				-0.284	0.753	0.566	-0.303	0.739	0.539

month19-Sep		-0.264	0.768	0.590	-0.276	0.759	0.573
month19-Oct		-0.107	0.899	0.795	-0.122	0.886	0.767
month19-Nov		0.014	1.014	0.971	-0.002	0.998	0.996
month19-Dec		0.488	1.629	0.246	0.476	1.610	0.255
month20-Jan		0.803	2.231	0.054	0.793	2.210	0.056
month20-Feb		0.897	2.452	0.038	0.887	2.428	0.039
month20-Mar		0.110	1.116	0.789	0.102	1.107	0.801
month20-Jul		-0.897	0.408	0.075	-0.919	0.399	0.068
month20-Aug		-0.982	0.375	0.051	-0.999	0.368	0.047
month20-Sep		0.318	1.374	0.573	0.292	1.339	0.602
AIC, BIC							
AIC		4595			4587		
BIC		4780			4747		

Table 5. Environmental GLMMNB Multivariate Model

	Coefficient	RR	p-value
Percent of landcover around household made up of trees	0.052	1.053	0.246
Distance to the nearest water body	-0.001	0.999	0.259
Average Temperature per month in Celsius	0.152	1.165	0.000
Percent of landcover around household made up of grass/shrubbery	0.000	1.000	0.995
Percent of landcover around household made up of dense vegetation	0.036	1.037	0.111
Fixed Effect			
Intercept	-1.352	0.259	0.263
Month (1-12)	-0.078	0.925	0.023
Random Effects			
Variance	0.671		

Std dev	0.819		
Dispersion Model			
(Intercept)	-1.656	0.191	0.000
month18-Nov	-0.140	0.870	0.694
month18-Dec	-0.341	0.711	0.357
month19-Jan	1.192	3.294	0.001
month19-Feb	1.367	3.923	0.000
month19-Mar	1.409	4.093	0.000
month19-Apr	1.666	5.290	0.000
month19-May	1.333	3.793	0.001
month19-Jun	0.665	1.945	0.142
month19-Jul	-0.765	0.466	0.134
month19-Aug	0.392	1.479	0.388
month19-Sep	0.193	1.213	0.661
month19-Oct	-0.381	0.683	0.273
month19-Dec	0.414	1.512	0.223
month20-Jan	1.027	2.792	0.003
month20-Feb	1.472	4.356	0.000
month20-Mar	0.566	1.761	0.102
month20-Jul	0.055	1.056	0.915
month20-Aug	-0.331	0.718	0.472
month20-Sep	0.651	1.918	0.097
Model Fit			
AIC	5056		
BIC	5195		

Table 6. Intervention GLMMNB Multivariate Model			
	Coefficient	RR	p-value
Time of outdoor gathering	-0.153	0.858	0.014
Number of nets hanging	0.233	1.263	0.004
Number sleeping in house	0.039	1.039	0.284
House sprayed within 6 months	0.004	1.004	0.405
Fire burning near indoor trap	0.011	1.011	0.956
Time of first exit in the morning	-0.036	0.965	0.679
Fixed Effect			
	5.338	208.106	0.062
	-0.022	0.978	0.564
Random Effects			
	0.724		
	0.851		
Dispersion Model			
(Intercept)	-1.207	0.299	0.000
month18-Nov	-0.329	0.719	0.413
month18-Dec	-0.647	0.524	0.120
month19-Jan	0.809	2.245	0.048
month19-Feb	0.881	2.413	0.032
month19-Mar	1.005	2.733	0.013
month19-Apr	1.159	3.185	0.005
month19-May	0.520	1.682	0.211
month19-Jun	-0.511	0.600	0.249
month19-Jul	-1.707	0.181	0.001
month19-Aug	-0.485	0.616	0.314
month19-Sep	-0.477	0.620	0.318
month19-Oct	-0.233	0.792	0.564
month19-Nov	-0.050	0.951	0.897

month19-Dec	0.324	1.383	0.415
month20-Jan	0.717	2.049	0.073
month20-Feb	1.068	2.911	0.011
month20-Mar	0.257	1.292	0.512
month20-Jul	-1.036	0.355	0.033
month20-Aug	-1.139	0.320	0.020
month20-Sep	0.185	1.203	0.732
AIC, BIC			
AIC	4839.000		
BIC	4987.000		

Figures:

Figure 1: Number of mosquitoes collected per month by subfamily

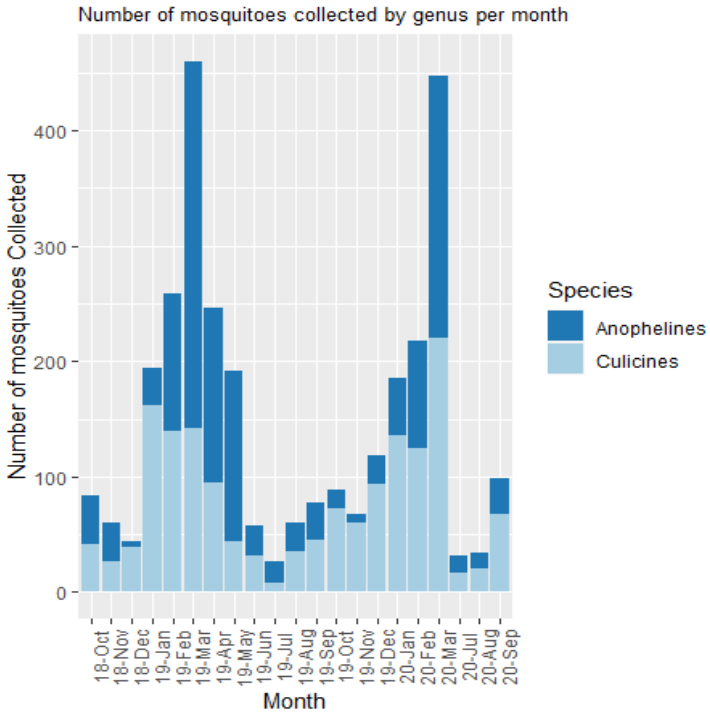


Figure 2. Plot to show the relative frequency of vector species tiers across season. '1' represents primary vector species, '2' is secondary, '3' tertiary, and '4' is non-malaria vector species.

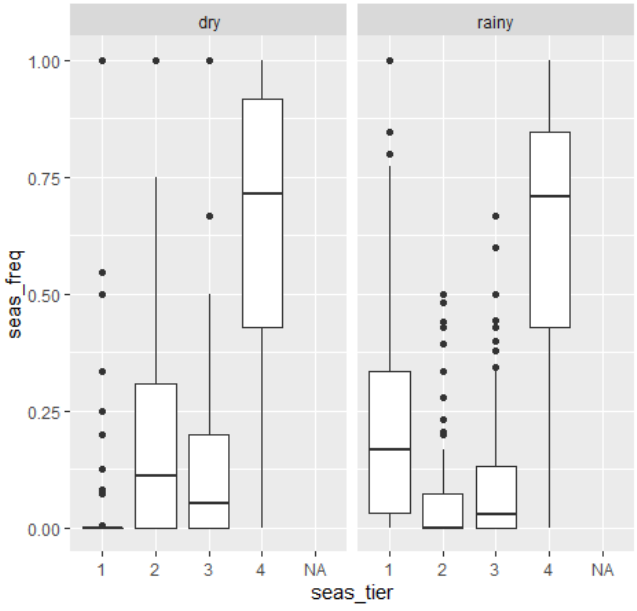


Figure 3. Relative quantities of blood-fed anophelines by trap location.

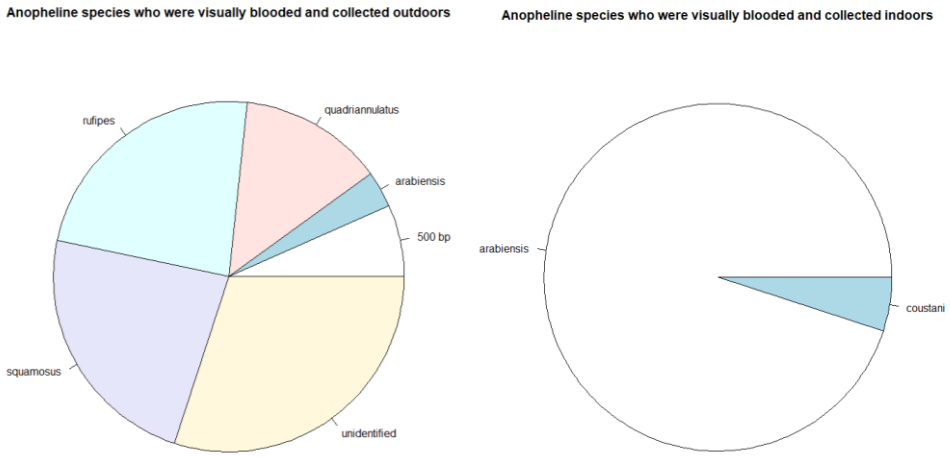
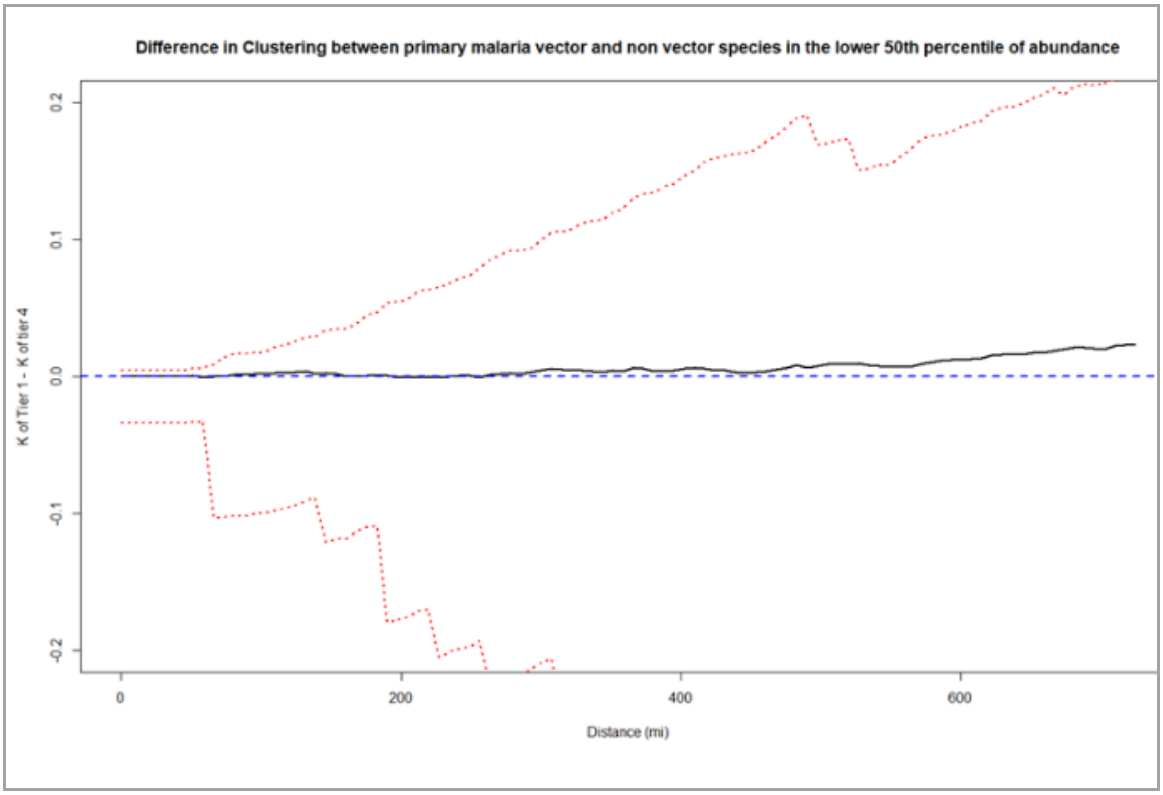
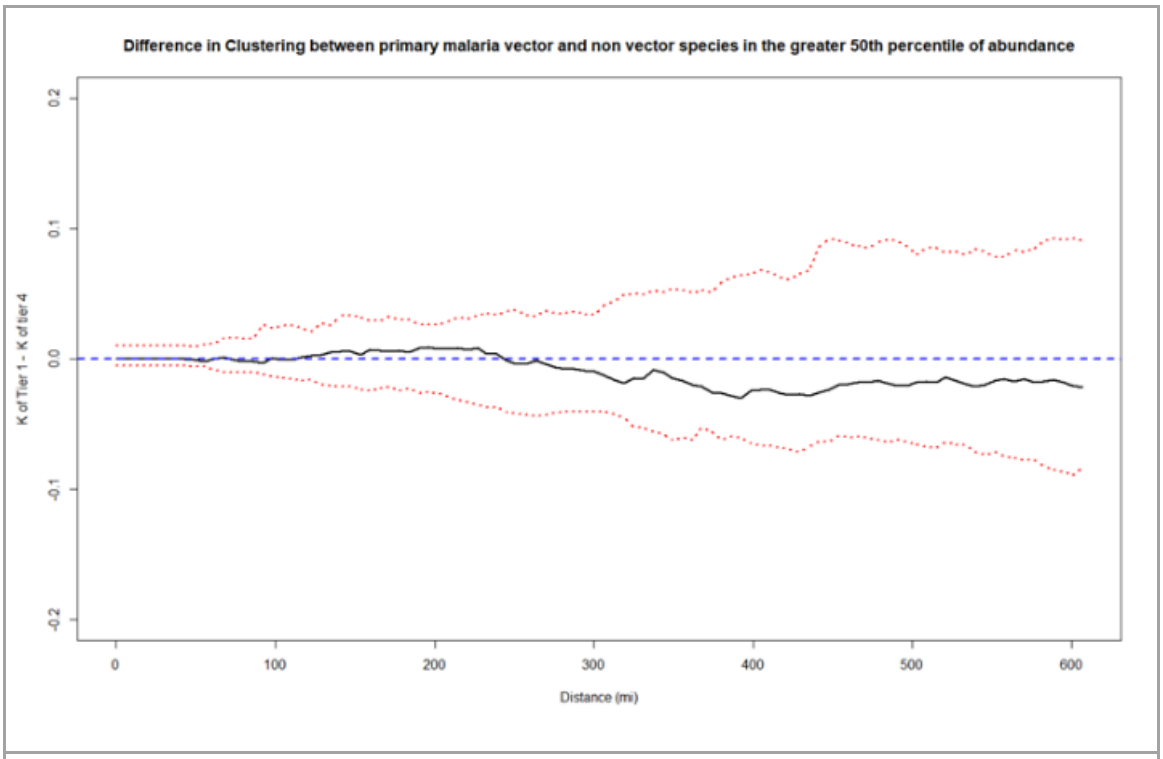
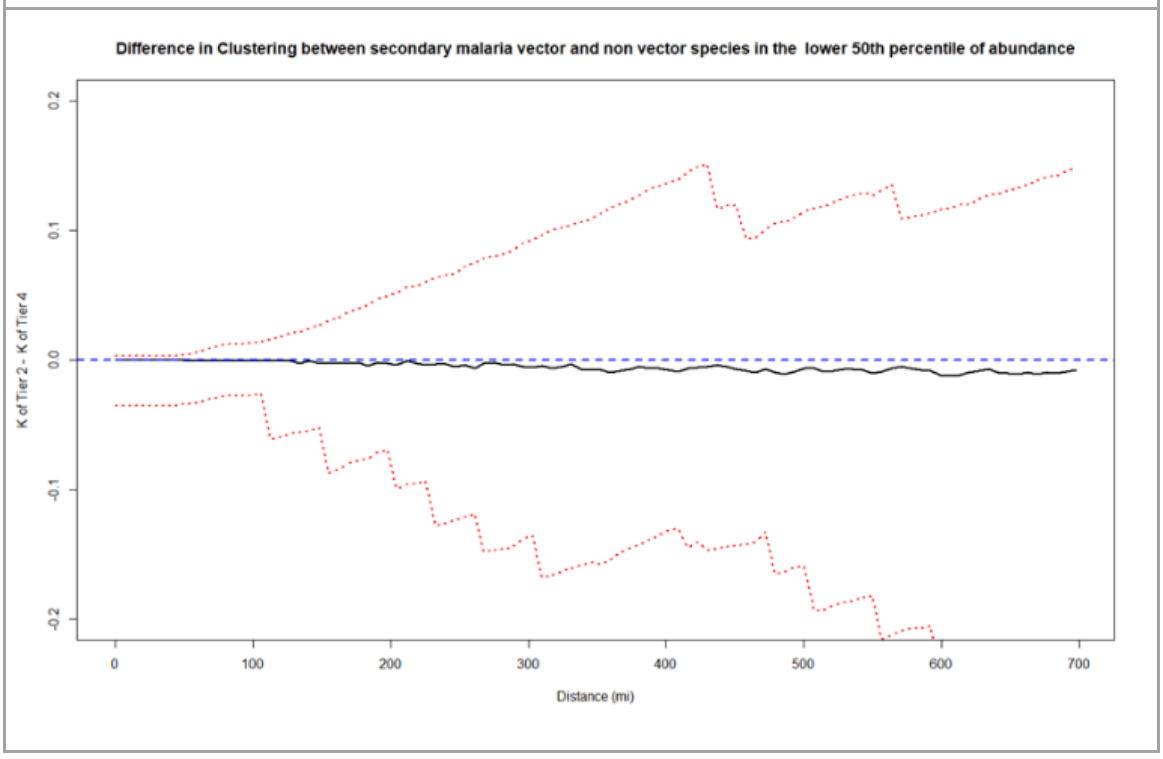
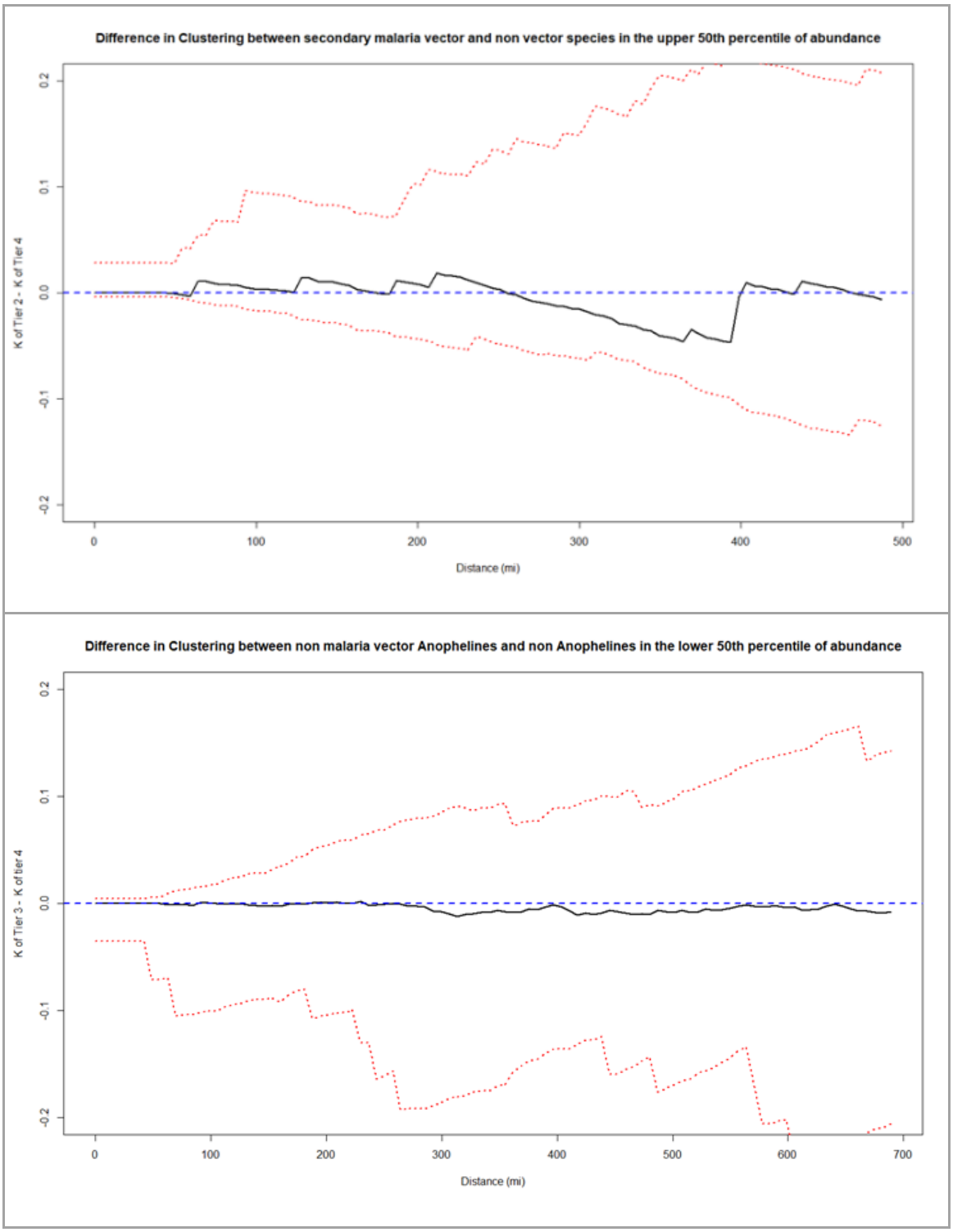


Figure 4. Difference in K-function plots between malaria-vector categories a non-vector species at high and low abundances.




Difference in Clustering between primary malaria vector and non vector species in the greater 50th percentile of abundance
K of Tier 1 - K of tier 4
0.2
0.1
0.0
-0.1
-0.2
0
100
200
300
400
500
600
Distance (mi)


Difference in Clustering between secondary malaria vector and non vector species in the lower 50th percentile of abundance
K of Tier 2 - K of Tier 4
0.2
0.1
0.0
-0.1
-0.2
0
100
200
300
400
500
600
700
Distance (mi)


Difference in Clustering between secondary malaria vector and non vector species in the upper 50th percentile of abundance
K of Tier 2 - K of Tier 4
0.2
0.1
0.0
-0.1
-0.2
0
100
200
300
400
500
Distance (mi)
Difference in Clustering between non malaria vector Anophelines and non Anophelines in the lower 50th percentile of abundance
K of Tier 3 - K of tier 4
0.2
0.1
0.0
-0.1
-0.2
0
100
200
300
400
500
600
700
Distance (mi)

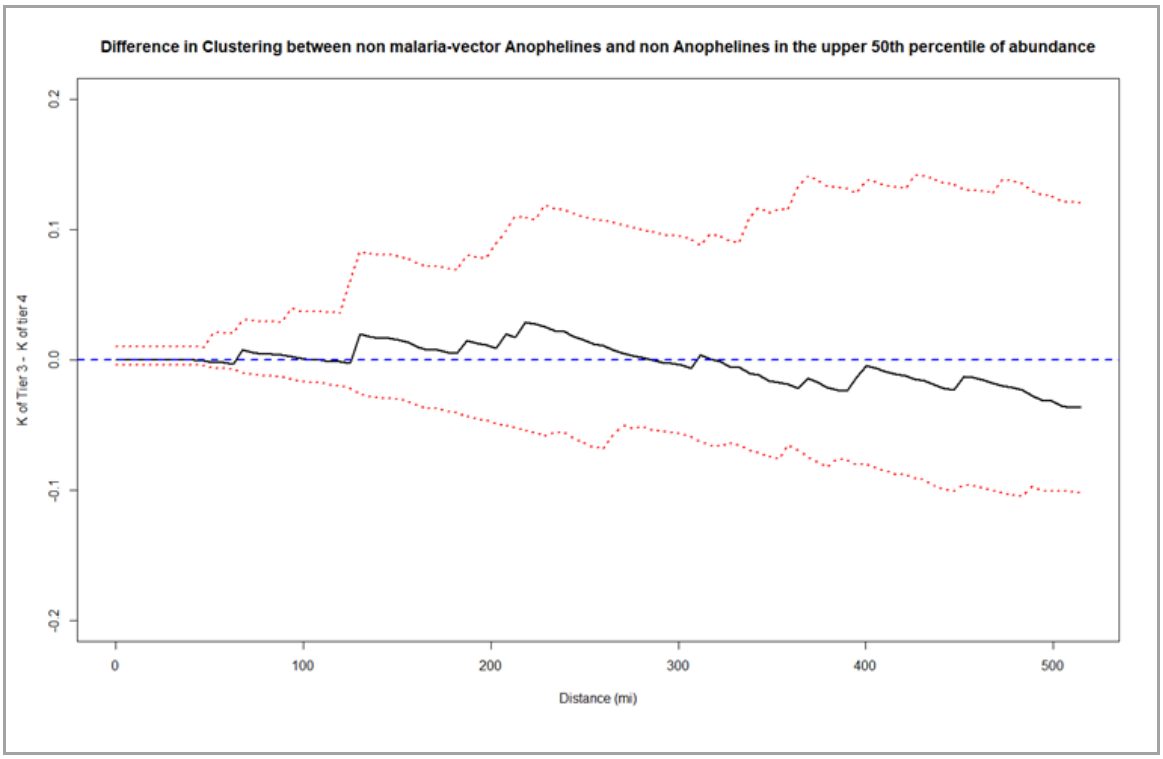
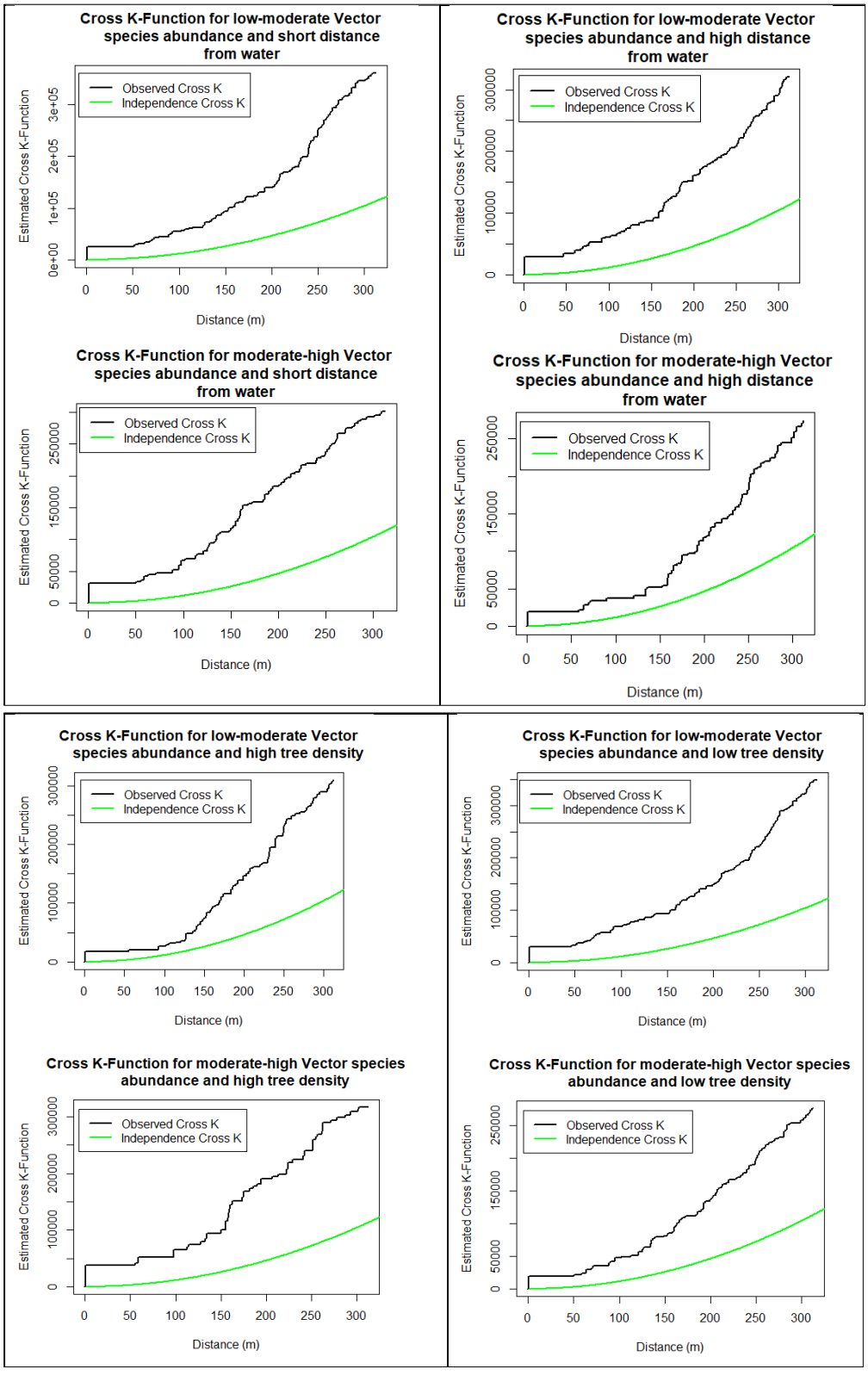
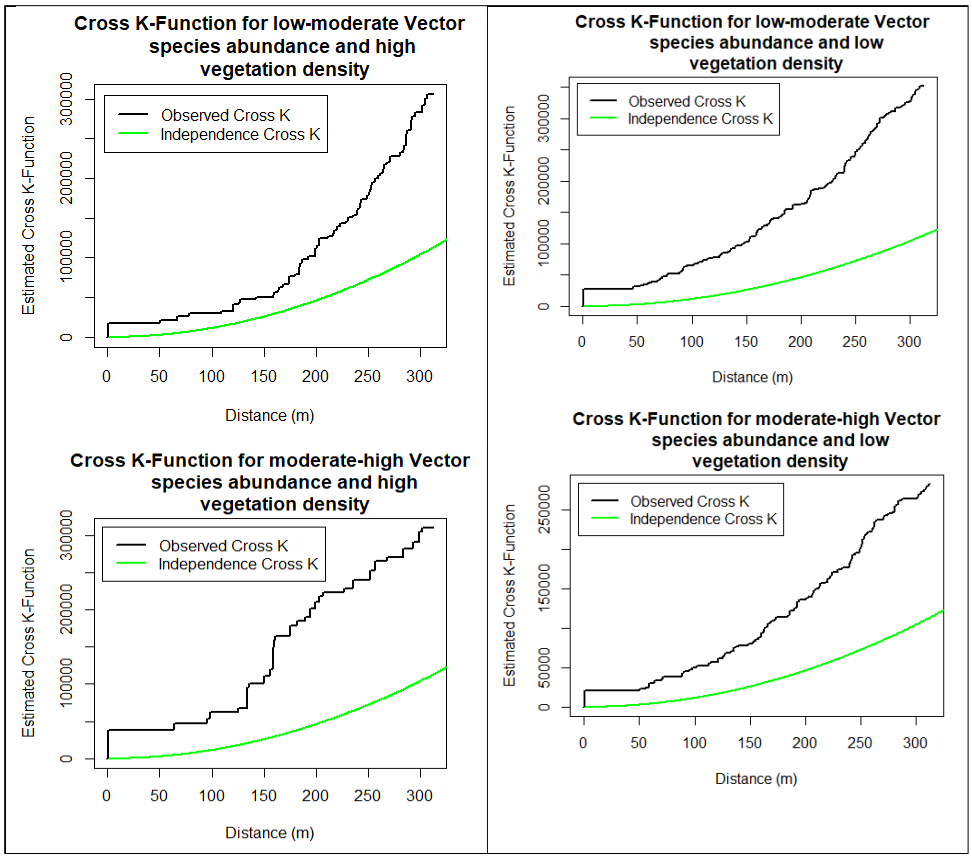

Difference in Clustering between non malaria-vector Anophelines and non Anophelines in the upper 50th percentile of abundance
K of Tier 3 - K of tier 4
0.2
0.1
0.0
-0.1
-0.2
0
100
200
300
400
500
Distance (mi)

Figure 5. Cross-K plots of *An. arabiensis* counts of different abundance and relative environmental variable densities around the household.




Cross K-Function for low-moderate Vector species abundance and high vegetation density
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)
Cross K-Function for low-moderate Vector species abundance and low vegetation density
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)
Cross K-Function for moderate-high Vector species abundance and high vegetation density
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)
Cross K-Function for moderate-high Vector species abundance and low vegetation density
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)

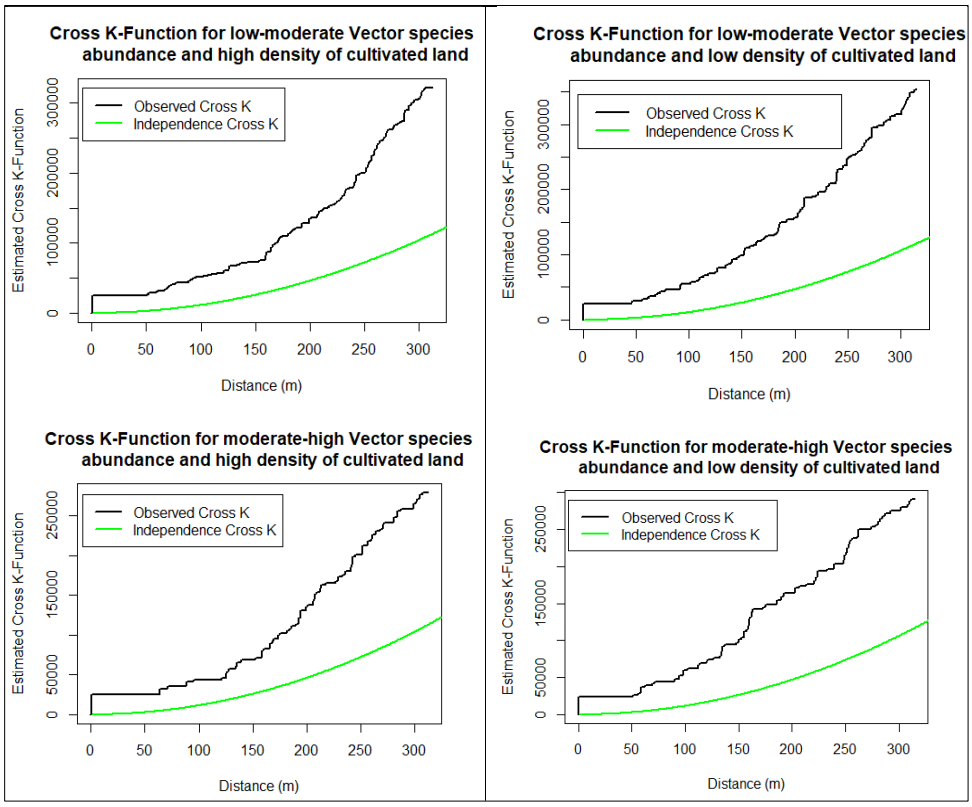

Cross K-Function for low-moderate Vector species abundance and high density of cultivated land
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)
Cross K-Function for low-moderate Vector species abundance and low density of cultivated land
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)
Cross K-Function for moderate-high Vector species abundance and high density of cultivated land
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)
Cross K-Function for moderate-high Vector species abundance and low density of cultivated land
Observed Cross K
Independence Cross K
Estimated Cross K-Function
Distance (m)

Figure 6. Morans I test of Multivariate Negative Binomial GLMM model 1 for distance based autocorrelation.

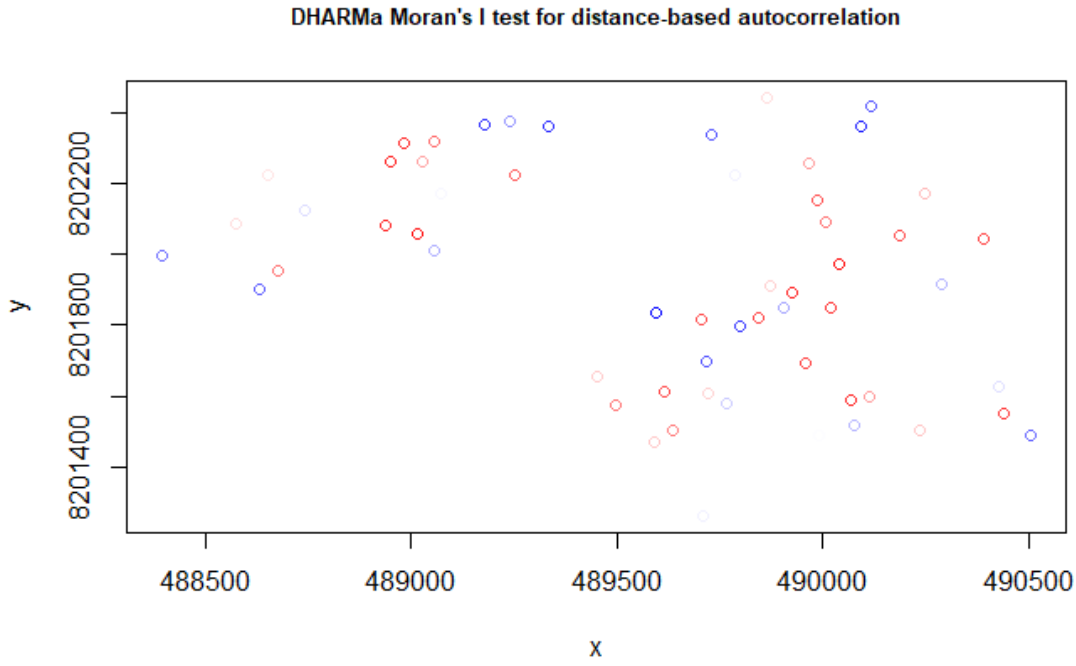
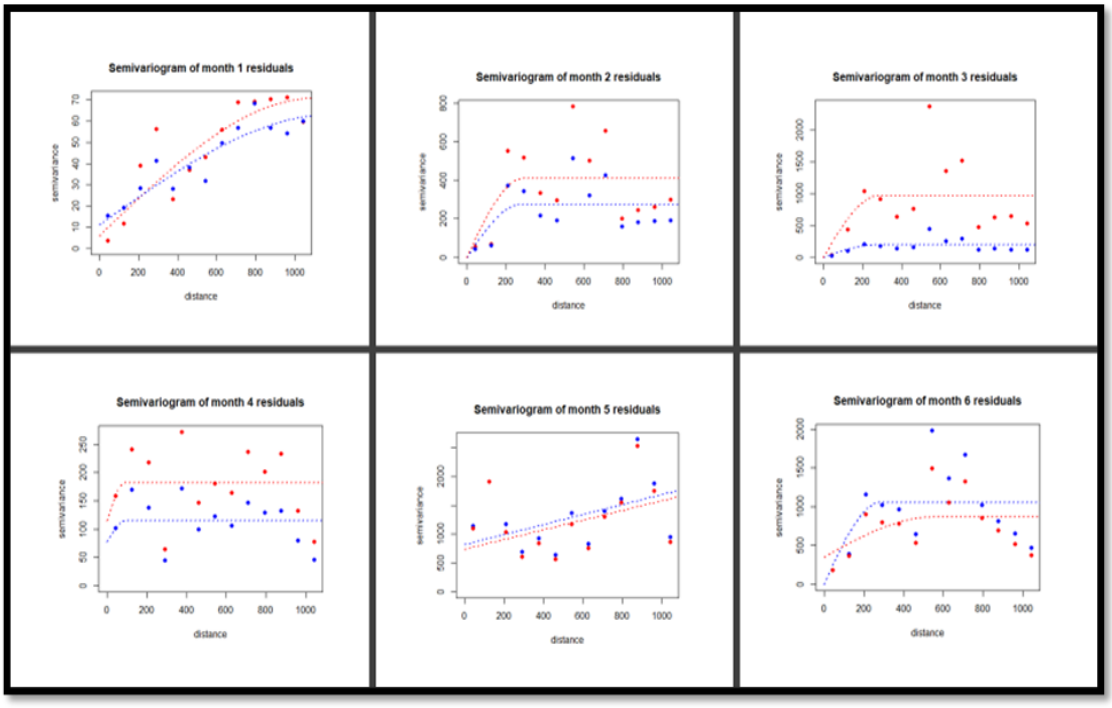


Figure 7. Semivariogram plots of residuals comparing an intercept only model (blue) with the adjusted model (red).



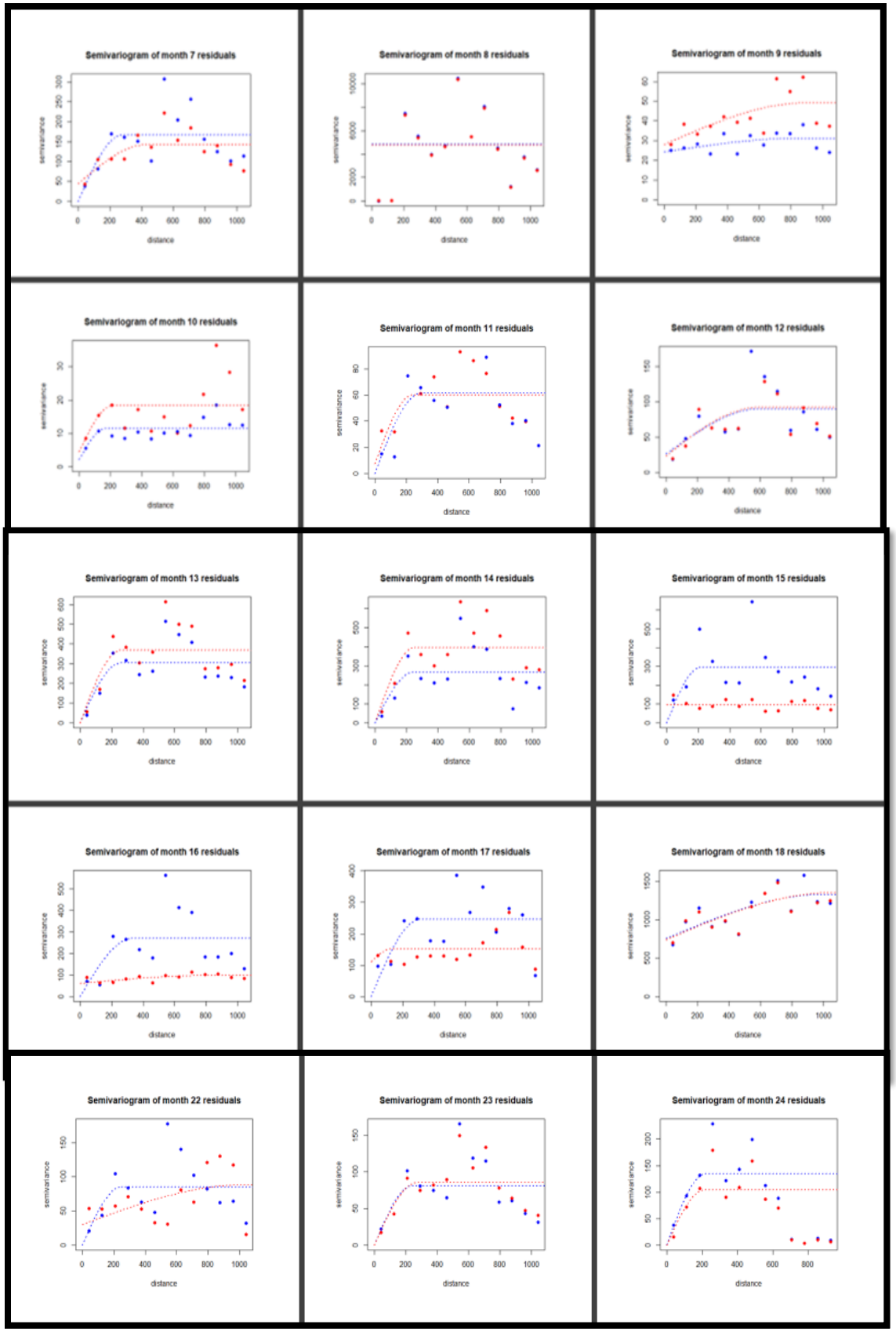

Semivariogram of month 7 residuals
Semivariogram of month 8 residuals
Semivariogram of month 9 residuals
Semivariogram of month 10 residuals
Semivariogram of month 11 residuals
Semivariogram of month 12 residuals
Semivariogram of month 13 residuals
Semivariogram of month 14 residuals
Semivariogram of month 15 residuals
Semivariogram of month 16 residuals
Semivariogram of month 17 residuals
Semivariogram of month 18 residuals
Semivariogram of month 22 residuals
Semivariogram of month 23 residuals
Semivariogram of month 24 residuals
semivariance
distance

Figure 8. DHARMa diagnostic tests of Multivariate Negative Binomial GLMM model 1.