Distribution Agreement

Spatial Distribution of Malaria Prevalence in Ethiopia Based on Village Level Environmental, Socio-Economic, and Demographic Parameters

By

Marc Cunningham
MPH 2011
Epidemiology

Uriel Kitron, MPH, PhD
Emory University
Faculty Adviser

Patricia Graves, MSPH, PhD
The Carter Center
Field Adviser

Spatial Distribution of Malaria Prevalence in Ethiopia Based on Village Level Environmental, Socio-Economic, and Demographic Parameters

By

Marc Cunningham
MPH 2011
Epidemiology

BA Plan II Honors
University of Texas at Austin
2007

BA Mathematics
University of Texas at Austin
2007

Faculty Advisor: Uriel Kitron, MPH, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2011

Abstract

Spatial Distribution of Malaria Prevalence in Ethiopia Based on Village Level Environmental, Socio-Economic, and Demographic Parameters

By Marc Cunningham

Environmental parameters detectable through remote sensing often correlate well with aspects of vector habitat or suitability. Investigation of these parameters in the context of existing malaria interventions provides a clearer understanding of the disease distribution. Specifically, we discovered that the previous two weeks average land surface temperature, the previous six months average daily rainfall, the previous two weeks average vegetation, and population density were significantly associated (positively or negatively) with prevalence of malaria in Ethiopia in 2006.

Stratification of risk in Ethiopia has historically been done based on elevation. Although altitude serves as a proxy for temperature, we suggest that temperature itself is the driving risk factor (Prevalence Ratio: 1.13). Further, given the role we saw these parameters play in disease distribution, we suggest that evaluations of programmatic work which fail to account for variations in environment over time may be biased.

Based on spatial clustering in the eastern portion of the State of Southern Nations, Nationalities and Peoples (SNNP) Region, and poisson regression results, we suggest further research into the role higher rural population density or peri-urban environments may play in focal transmission.

Spatial Distribution of Malaria Prevalence in Ethiopia Based on Village Level Environmental, Socio-Economic, and Demographic Parameters

By

Marc Cunningham
MPH 2011
Epidemiology

BA Plan II Honors
University of Texas at Austin
2007

BA Mathematics
University of Texas at Austin
2007

Faculty Advisor: Uriel Kitron, MPH, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2011

Acronyms

AIC............. Akaike information criterion

CCD .... Cold Cloud Duration

CIESIN....... Center for International Earth Science Information Network

EVI .... Enhanced Vegetation Index

FMOH .... Federal Democratic Republic of Ethiopia Ministry of Health

GFATM .... Global Fund to Fight AIDS, Tuberculosis and Malaria

GIS .... Geographic Information System

GPW .... Gridded Population of the World

IRI............... International Research Institute for Climate and Society

IRS.............. Indoor Residual Spraying

LISA .... Local Indicator of Spatial Association

LLIN .... Long Lasting Insecticide Treated Net

LST .... Land Surface Temperature

MIS............. Malaria Indicator Survey

NDVI .... Normalized Difference Vegetation Index

NDWI .... Normalized Difference Water Index

NOAA .... National Oceanic and Atmospheric Administration

SEDAC ...... Socio-Economic Data and Applications Center

SNNP......... State of Southern Nations, Nationalities and Peoples

TCC............ The Carter Center

USGS.......... United States Geological Survey

Acknowledgments

I would like to thank my advisors. I sincerely appreciate Patricia Graves for her help in selecting the topic, for introducing me to The Carter Center, to her collaborators, and to this data set. She provided valuable comments, edits, and encouragement. I am grateful to Uriel Kitron for taking on an epidemiology thesis student, for his insight into the roles ecology and biology play in spatial analysis, for his high expectations, and for his unwillingness to let me be satisfied with a lesser thesis. I look forward to working with you both as we move forward to publication.

Many thanks to Deb MacFarland, for introducing me to Tricia and The Carter Center, and for bringing me news from Nigeria. It was a privilege to meet Jeremiah Ngondi, who provided perspective on the 2006 and 2007 survey designs and implementation, and on prior analysis. Pietro Ceccato and Madeleine Thomson from IRI provided support in acquisition and interpretation of the environmental data. I wouldn't have had much to work with without you.

I am indebted to Julie Clennon, Lance Waller, Gonzalo Vazquez-Prokopec and Donal Bisanzio for their roles in introducing me to the companion fields of Geographic Information Systems and Spatial Epidemiology. Julie and Lance provided ongoing support even after I finished their class.

I am grateful for Susan Allen, who has served as my boss, mentor, and friend through my time as a masters student here at Rollins School of Public Health.

Paul Weiss's insight into both SAS and complex survey design continue to astound me.

To my cohort––thank you for dragging me out of the library or away from my table to remind me that there's fun to be had in Atlanta. I thank the Epi Department staff and faculty for their classes, their open doors, and for many conversations on nuts and bolts of Epi.

We acknowledge The Carter Center's and the Ethiopia Ministry of Health's role in data collection and program implementation, and thank all participants for donating their time during the surveys. ––*Marc Cunningham*

Table of Contents

Introduction to Spatial Epidemiology

This article discusses the spatial distribution of malaria caused by *Plasmodium vivax and P. falciparum* in Ethiopia. For those unfamiliar with spatial epidemiology, we first outline some tools, techniques, and perspectives of the field.

Spatial epidemiology is concerned with understanding spatial determinants of transmission and risk, typically for the purposes of better disease control. It often combines case data from field research or surveillance with data on environmental and ecological factors by overlaying them in a georeferenced database. It compares the observed distributions with what would have been expected if there were no spatial clustering and if there were no association between the factors of interest and the disease.

Tobler's first law of geography states, "Everything is related to everything else, but near things are more related than distant things" [1,2]. Based on this axiom, spatial epidemiology aims to answer the following questions [3,4]:

(1) Is there clustering of disease or of intensity of disease transmission?

(2) If so, where are these 'hot spots' (of high disease) and 'cold spots' (of low disease)?

(3) What is unique or similar about these areas that explains the clustering? and

(4) How does this knowledge translate to improved control and prevention programmes through prediction of risk?

A few common tools for evaluating clustering are Moran's I and K-Functions (global level clustering) and Local Indicators of Spatial Association (LISA) and Getis Gi*(d) (local level clustering). Moran's I quantifies the correlation between similarity of outcomes based on proximity [5] [6] [7]. The K-function accounts for intensity of disease in cluster analysis and provides insight into the distances at which clustering is occurring [8,9]. The LISA detects localized clustering by comparing the similarity of disease intensity around each outcome [10]. Getis Gi*(d) estimates local correlation in a defined area around (but not including) each outcome, and compares that correlation with the global correlation to find areas of high or low clustering [11,12]. As it creates these comparisons for each outcome, the test statistic generated must be adjusted to account for multiple comparisons.

Vector borne infectious disease patterns, specifically, are driven by both intervention programmes and by environmental parameters [3,13,14]. Rainfall, temperature, vegetation, elevation, slope, aspect, soil type, soil texture, soil moisture, hydrology, water bodies, anthropogenic biomes and/or landcover can often be linked with aspects of vector habitat and suitability, or parasite development. In affecting vector habitat, these parameters shape vector distribution, which affects the interaction between vector and the human populations, and the resulting case distribution.

Satellite remote sensing provides large scale datasets on these parameters [15,16]. Imagery has spatial, temporal, and spectral resolution. In passive sensors, the satellite sensors absorb electromagnetic (EM) radiation that has been reflected or emitted from the earth. Spectral resolution refers to the range of the EM radiation that the sensor observes. High spectral resolution allows better discrimination of landcover, as the contrast between reflected EM radiation in different bands provides more information on environment than a single band. Spatial resolution refers to the amount of area from which EM radiation is averaged to create a pixel. Finer resolution allows better discrimination of microenvironments. Temporal resolution refers to the frequency with which the satellite(s) image the area. Several high temporal resolution images are often combined into composites, which provide more stable and accurate parameter estimates over the time period of interest. Other sources of environmental data, such as weather stations and field collection, can be used to improve or validate remote sensed data. Though good for specific locations, these are generally more costly, and are often unfeasible for large areas.

Remote sensed indices are created by combining and contrasting different bands of EM radiation based on the reflective properties of the environmental parameter of interest. Briefly, four commonly calculated remote sensed indices are the Normalized Difference Vegetation Index (NDVI) [17] which provides information on vegetation biomass and coverage, the Normalized Difference Water Index (NDWI) [18] which provides information on vegetation health by measuring vegetation water content, Land Surface Temperature (LST) [17,19], and Cold Cloud Duration (CCD) [15] which estimates rainfall. A fifth common remote sensed parameter is elevation. A Digital Elevation Map (DEM) created by the NASA Shuttle Radar Topography Mission (SRTM) [20] provides high spatial resolution global estimates.

Satellite imagery has become increasing available at finer temporal and spatial scales. Caution is needed, however. These environmental data cannot replace high quality case information, or field ecology work on the vectors [21,22]. Habitat suitability does not necessarily correspond with vector density and/or disease presence, nor with opportunities for vector and susceptible human host interaction [23,24].

Geographic Information Systems (GIS) allow processing, correcting, and overlaying of geographically coordinated data [15]. Satellite data must be pre-processed before being analyzed to remove contamination [15]. Imagery must then be projected, converting a mostly spherical world onto a flat map. The resulting clean, projected environmental parameters can be related to georeferenced case information.

Associations between risk factors and disease can be examined by overlaying maps, and quantitatively through regression [25,26,27]. Regression models can fit a variety of outcome distributions, such as: normal [28], binary [29,30], poisson [31] [32], negative binomial [33], and zero-inflated poisson [33]. These models are constructed by finding the best fit line, often through maximum likelihood estimation techniques. One assumption in typical regression modeling is that the predictors and outcomes are independent between different individuals. As with time-correlated data, spatially correlated data fail to meet this assumption. The inclusion of additional spatially dependent regression terms can account for this correlation, allowing better insight into both global and local trends [6,34,35,36].

When examining the spatial distribution of disease, the concept of scale plays a large role in framing research [37,38,39]. Differing temporal and spatial scales relate to different pathogen, individual, or population level characteristics [21]. Understanding at one scale does not always translate to finer or coarser scales [40]. From a pragmatic perspective, finer scales—temporal, spatial, or spectral—have higher costs, and often have a lower signal to noise ratio, making long term or regional patterns more difficult to observe [40,41]. At the cost of time and computing power, one can always move up scale from fine to coarser data by averaging appropriately.

The final goal of spatial epidemiology is an enriched understanding of the biological, sociological, or environmental parameters which underlay disease distribution and transmission [23,42]. Ideally, these parameters can be used in the formation of a risk map [4,21,43,44]. A more complete perspective on disease dynamics provides the foundation for more efficient and effective control and prevention programmes, and context for programme evaluation.

As in all models, there is a trade off in spatial models between generality, precision, and realism [45]. Often the better our model explains the minutia of a specific time and area, the worse it tends to perform in different contexts. A risk map, in particular, requires validation if extended beyond the original study area or time [37]. When used appropriately, however, risk maps provide an example of how spatial epidemiology can be translated into public health policy through targeted allocation of resources proportional to risks.

Background

Historical Malaria in Ethiopia

Malaria offers a significant challenge as a large cause of mortality and morbidity in Ethiopia. In 2009 alone, 3 million suspected cases sought health assistance [46]. There are estimates of 9 million cases each year with 70,000 deaths, with over 6 million additional cases and 114,000 deaths in epidemic years [47]. Despite recent gains in decreasing prevalence, around 68% of the population of the country's population of 78 million persons remain at risk based on estimates from altitude and the current regional extent of disease–– there is potential for resurgence [47,48]. Transmission is both seasonal—with most areas experiencing peak transmission between September and December[1]—and cyclical, with widespread epidemics every five to eight years [47]. Widespread epidemics have been recorded since the 1950s, including highland epidemics in 1988, 1991, 1998, and the most recent in 2003 [46,48].

Malaria and Ecology

Mosquitoes of the *Anopheles gambiae* complex are the most common mosquitoes in Ethiopia, [50], and are dominated by *An. arabiensis* [51], which breeds well in temporary, often shallow sunlit pools formed in pits, footprints, or tire tracks [52]. Other vectors of interest include *An. funestus*, *An. nili* and *An. pharoensis* [47]. Malaria, spread by the *Anopheles* sp., is caused by infection with the *Plasmodium* sp.. The dominant forms of

[1] Specific regions with a second season of rains experience transmission in April to May 49. FMOH (2006) National Five Year Malaria Control Plan .

malaria in Ethiopia are *P. falciparum* and *P. vivax*. The former is more common and accounts for roughly 70% of infections [48,53].

In Ethiopia, malaria existing control policies are based on the premise that ecologic suitability corresponds with altitude, with seasonal and annual temperatures, and with seasonal and annual rainfall. Transmission has historically been stratified into the following zones based on suitability and historical patterns: stable year round, seasonal, unstable, arid and dependant on water bodies, and free [47]. Due to the heterogeneous nature of malaria transmission in Ethiopia, locations of malaria transmission do not always match these historically mapped stratifications.

Temperature, habitat suitability, humidity, and rainfall interact to influence vector density, longevity, and behaviour, together with human actions which are based on perceptions of risk. Temperature also influences the rate of parasite development [54] [55]. These complex interactions affect host contact with infective mosquitoes, and the resulting distribution of malaria prevalence [56,57]. Satellite remote sensed nightly land surface temperature (LST) , Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and precipitation estimated from cold cloud coverage and weather stations have been shown to correlate with air temperature; vegetation, larval sites, and humidity; and rainfall [26,29,58,59].

Current Malaria in Ethiopia

With the receipt of funding from the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) in 2003 (Round 2: 2003 – 2008; $73 million) and in 2005 (Round 5: 2005 – 2010; $140 million) the Ethiopian Federal Ministry of Health (FMOH) scaled up of malaria prevention and control [47,48], including nationwide distribution of long-lasting insecticidal nets (LLINs) and artemether-lumefantrine treatments [49]. One strategic goal for the National Malaria Control Program is distribution of insecticide treated nets to 100% of households in malarious areas, with the corresponding objective of appropriate net use by at least 90% of pregnant women and children under 5 [49]. The Carter Center (TCC) has partnered with the FMOH in bednet distribution and education programmes since early 2006, and have assisted in the distribution of over three million nets in three targeted regions of Amhara, Oromiya, and the State of Southern Nations, Nationalities and Peoples (SNNP) [60]. The next Malaria Indicator Survey (MIS) is planned for later this year, and will provide updated estimates on programmatic success as measured through parasite prevalence, anemia, malaria knowledge, net ownership, net quality, and net use.

Survey Information

In coordination with the FMOH, The Carter Center conducted a 2006 household Baseline Survey (2006 Baseline) in Amhara, Oromia, and SNNP [61,62,63]; and the 2007 Ethiopia National Malaria Indicator Survey (2007 MIS) in collaboration with the Ethiopia MIS study group[2] [48,64]. These two surveys provide information on the relationship between demographics, behavioral risk factors, socio-economic status and malaria prevalence in Ethiopia. Survey methods for both surveys have been previously discussed [60,61,62,64,65]. The 2006 Baseline and 2007 MIS received ethical approval from the Emory

[2] Federal Ministry of Health of Ethiopia, The Carter Center, Malaria Control and Evaluation Partnership for Africa (a programme at PATH), World Health Organization, United Nations International Children Emergency Fund, U.S. Agency for International Development, U.S. Centers for Disease Control and Prevention, the Central Statistical Agency, Center for National Health Development in Ethiopia, and Malaria Consortium {Jima, 2010 #22}

University Institutional Review Board (IRB 1816, 6389), the Ethiopian Regional Health Bureaus, the PATH Ethical Committee, and the Ethiopian Science and Technology Agency.

Briefly, the 2006 Baseline took place from Dec 2006––Feb 2007 and was designed to estimate malaria prevalence, net ownership, and net use prior to the scale up in LLIN distribution. 5,708 households were sampled in 224 clusters for a total of 27884 participants [62]. Clusters were randomly selected from non-urban areas in districts where at least 90% of the population lived in malarious areas, as defined by expert knowledge and regional criteria [62]. Parasitology was evaluated in every other household, with prevalence and type of malaria determined through blood slide microscopy [61].

The 2007 MIS took place from Oct through Dec 2007 and was designed to assess progress toward national roll back malaria goals by investigating access, coverage, and use of malaria prevention and control interventions. 7,621 households were sampled in 319 clusters for a total of 32,380 participants [64]. All children under five were tested for malaria together with all household members in every fourth house. The three regions of interest were oversampled to provide power for comparisons with the 2006 Baseline. Based on microscopy (10,578 slides), the 2007 MIS survey estimates a national age-adjusted prevalence of 0.2%. Overall age adjusted prevalence decreased in the three regions of from 3.5% to 0.1% [60].

Prior analyses have described individual and household characteristics including indoor residual spraying, household net ownership, net usage, and malaria prevalence; and have related the latter to routine surveillance data [60,63]. Between 2006 and 2007 the ownership and overall use of LLINs significantly increased. In Amhara, Oromiya, and

SNNP, malaria prevalence significantly decreased in all age groups [60]. In the 2006 Baseline survey, altitude (aOR 0.95 per 100 m), the asset index (aOR 0.79 per quintile), rainfall (aOR 1.10 per 10mm) and the number of LLINs per house (aOR 0.60), were significantly associated ($p<0.05$) individual malaria prevalence [62]. In the 2007 MIS survey, use of LLINs was significantly associated with malaria knowledge (aOR 2.1), altitude (aOR 0.3), and IRS (aOR 4.6) [53].

Research Hypothesis

This research examines the spatial distribution of malaria in the three regions of Amhara, Oromiya, and SNNP with respect to ecological and environmental factors while controlling for demographics and LLIN coverage/LLIN use at the kebele (village) level. Having discussed the methodology and background for both surveys, we analyze the 2006 Baseline data.

We hypothesize that although socio-economic status (SES), and LLIN ownership and use are driving factors of malaria prevalence at the individual scale [62], environmental factors drive the spatial heterogeneity of malaria in Ethiopia at the village scale. We investigate temperature, rainfall, vegetation, hydrology, proximity to water, landcover, population density, and altitude. With access to cross-sectional information on malaria prevalence, the we focus on spatial rather than temporal variation.

Substantial research has been done looking into the complex relationship between prevalence of malaria, malaria's ecological niche, and the relationship between remote sensed data and that niche [19,23,51,58,59,66,67,68,69,70]. Less research has been done examining the interaction and relative importance of these parameters while considering fine scale demographic, behavioural, and programmatic factors. A recent article by Ashton et al. partially fills this gap, but is limited in spatial scale to Oromiya [71].

Methods

Spatial Analysis of Clustering

We tested global clustering using Moran's I based on inverse distance squared, considering presence or absence of malaria at the village level, and then considering prevalence. Following this, we used Local Indicators of Spatial Association (LISA) based on presence or absence and then based on prevalence to determine specific hot-spots and cold-spots, with a test statistic of $z > 1.96$ considered significant. In considering Moran's I and LISA, we first examined the entire study area, and then examined Amhara alone compared with SNNP and Oromiya––we aggregated the regions based on the difference in times sampled.

Regression Variables

ArcGIS was used to match kebeles georeference with GPS units to environmental covariates.

As the prevalence of malaria across kebeles was non-normal, we used weighted poisson regression, modeling the number of prevalent cases sampled per village as our outcome with an offset of the number of people sampled. Ordinal logistic regression was considered, and rejected based on a failure to satisfy the proportional odds assumption.

All individual level characteristics––age, sex, and use of any net or a LLIN the night before the survey––were averaged to the village level, accounting for selection probabilities.

Household characteristics aggregated to the kebele level include a household socio-economic status (SES) index, ownership of at least one net, ownership of at least one LLIN, number of nets and of LLINs, a ratio of the number of nets and of the number of LLINs to the household size, and indoor residual spraying (IRS) within the last 6 months or within the last year. Selection probabilities were summed to the kebele level to provide weights.

At the kebele level we considered average altitude, average daily rainfall, average nighttime land surface temperature (LST), proximity to lakes, rivers, or hydrolines which can serve as a proxy for irrigation and mosquito breeding habitat, average vegetation which can serve as a proxy for humidity, anthropogenic biomes [72], and human population density [73] (See Table 1).

The 2006 Baseline survey took place from 12 Dec 2006 to 07 Feb 2007. Sampling of a kebele took 1 to 2 days. We divided the survey into two groups based on the median time-period (31 Dec 2006). In linking environmental factors to surveys, we considered all surveys in the first half to have been collected on 22 Dec 2006 (Period 1) and in the second half to have been collected on 18 Jan 2007 (Period 2). These time periods were chosen as balance between the Period midpoint and the availability of remote sensed data. All surveys in Period 1 were in Amhara, while Period 2 surveys included the eastern edge of Amhara, together with the SNNP and Oromiya regions. We examined all environmental parameter means, and then stratified the means by time period sampled. (See Table 1).

For LST, vegetation and rainfall we considered various aggregations in time. For LST we looked at a continuous range of degrees Kelvin (K) from composites of 1 to 16 days before the survey period (LST1), 17 to 32 days before (LST2), 1 to 32 days before (LST3), and

1 to 48 days before (LST4), all of which were averaged over a 1500m radius from the kebele centroid. Temperature estimates were acquired from USGS LandDAAC Modis with a 1km spatial resolution, and 8 day composite temporal resolution. The 1500 meter radius was chosen based on mosquito flight range estimates. To examine vegetation we considered the NDVI and NDWI values over the same distance and time periods as LST. Vegetation indexes were acquired from USGS MODIS with 16 day composites and 250 meter spatial resolution [17]. Rainfall was estimated from the National Oceanic and Atmospheric Administration (NOAA) Famine Early Warning System at 11 km spatial resolution and ten day composites [74]. We considered one month (NOAA1), two months (NOAA2), three months (NOAA3), six months (NOAA4), and nine months (NOAA5) prior to the survey period, extracted directly from the kebele centroids. The International Research Institute for Climate and Society[3] (IRI) provided the interface used to download these environmental parameters at the above mentioned temporal scales.

We estimated the kebele altitude as the average altitude for all households sampled in the kebele, which was similar to that extracted from the kebele centroid from the NASA Shuttle Radar Topographic Mission (SRTM) Digital Elevation Map (DEM) [20].

Proximity to water bodies (rivers and lakes), and proximity hydrolines were based on the ERSI Maps v10 with buffers of 500 meters, 1000 meters, 1500, and 3000 meters. Landcover [75] and population density [73] were acquired from the Socioeconomic Data and Applications Center (SEDAC) at the Center for International Earth Science Information Network[4] (CIESIN). We *a priori* reclassified anthropogenic biomes from 20 classes [75] to

---

[3] http://portal.iri.columbia.edu
[4] http://www.ciesin.columbia.edu

five: dense settlements, croplands, rangelands, forested areas, and wildlands. For the purposes of stable estimates, we required at least five kebeles within an environmental buffer to include it in analysis. Thus, proximity to water bodies was excluded from analysis.

Manipulation of geospatial data was done in ArcGIS 9.2 (ERSI). Regression was performed using SAS 9.2 (SAS)

Modeling strategy:

Based on the non-normal distribution of the village prevalence values, we used a poisson regression model. Over-dispersion was present (Pearson Deviance / DF ~ 3.8), which can lead to inappropriately tight confidence intervals. We scaled the residuals by the square root of the Pearson Deviance divided by its degrees of freedom.

Univariate and multivariate poisson regression was done in SAS using the weight statement to account for kebele sampling probabilities. Sampling stratification by region was not accounted for, resulting in conservative confidence intervals. Given that we aggregated all data to the primary sampling unit (kebele), it was not necessary to account for complex survey design in analysis.

We constructed models using a combination of discriminant analysis and stepwise regression. We first used the Akaike information criterion (AIC) [76] to select the best environmental variable from each group of temperature, vegetation, and rainfall—we considered all possible models with (or without) elevation, reclassified anthropogenic biomes and at most one temperature, vegetation, rainfall time period. The model with the lowest AIC was selected as the base environmental model. Further modeling was done with stepwise regression, reconsidering environmental variables at each step. We *a priori* selected $p<0.15$ as criteria for entry, and $p>0.20$ as criteria for removal.

Multi-collinearity was assessed for all variables in the final model. To avoid multi-collinearity and time-correlated covariates, only one rainfall, one vegetation, and one temperature time frame were included in the final model. Further, with LST in the model, altitude was not included.

Results

Spatial Analysis

When considering the entire study region, no clustering was observed. At a finer scale when considering only Oromiya and SNNP, there was significant (p=0.01) global clustering for the presence/absence of malaria at the kebele level.

Based on LISA for the entire study area, we saw only one area of Hi-Hi clustering, on the border of Amhara and Oromiya. When considering the north and south separately (Amhara alone, Oromiya with SNNP), a larger local cluster in eastern SNNP was apparent.

Table 0: Moran's I Evaluating Global and Regional Clustering of Malaria in Amhara, Oromiya, and SNNP (2006 Baseline)

	Presence/Absence		Prevalence	
Region	**Moran's I**	**P-value**	**Moran's I**	**P-value**
All	-0.13	0.54	-0.13	0.53
Amhara	-0.18	0.43	-0.13	0.54
Oromiya and SNNP	0.54	0.01	0.33	0.08

Descriptive Results

Across the study area [60], the average prevalence was 4.1%, varying from 0.9% in Oromiya, 4.6% in Amhara, and 5.4% in SNNP. Prevalence differed significantly between Period 1 at 5.4% (95% CI 4.0 – 6.1), and Period 2 at 2.0% (95% CI 1.6 – 3.3). 27.8% of the study population used a net the night before the survey, while 15.3% used a LLIN. The average number of LLINs per person was 0.3. Kebele net coverage was assessed as the percentage of households owning at least one LLIN, and was significantly protective (p=0.04) in our final model. Regarding the extent of coverage: 37.0% of households owned at least one bednet, and 21.5% owned at least one LLIN.

In all time lags considered, the mean of kebele temperatures was similar per period. There was more variability amongst kebeles in Period 2 (Std Dev 0.19-0.20) compared with Period 1 (Std Dev 0.28). NDVI and NDWI correlated strongly with each other, and average NDVI's and NDWI's associations with prevalence were similar in direction and significance per period across all time lags. For Period 1, rainfall averaged over the six months prior to the survey (NOAA4) was highest, while for Period 2 rainfall over the 9 months prior to the survey (NOAA5) was highest, reflecting differences in rainfall patterns and survey timing. On average, rainfall in the first period was higher than in the second. In the month prior to the survey, however, average daily rainfall (NOAA1) was slightly higher in the second period at 7.6 mm (95% CI 6.4 - 8.9) than in the first at 8.6 mm (6.3, 10.8).

To recap: average LST was similar between the sampling periods, though it varied greatly by kebele. Average rainfall, in comparison, differed by Period for NOAA3, NOAA4, and NOAA5, with Period 2 receiving less rainfall. Average rainfall in the month prior to the survey, which was marginally associated in univariate analysis with a decrease in prevalence, was lower for kebeles in Period 1 (7.6 mm) than in Period 2 (9.2 mm).

Univariate Associations in Poisson Regression

In univariate analysis at the village scale, population density (Prevalence Ration [PR] 1.16 per 100 people per km^2, p=0.03), LST3 (PR 1.08 per 1 K, p=0.05), and LST4 (PR 1.09 per 1 K, p=0.03) were significantly associated with increased prevalence. Non-significant associations were observed between prevalence and proximity (within 3000 meters) to water (PR 1.45, p=0.16); and between prevalence and vegetation in the 16 days prior to the survey (NDVI1 PR 0.44, p=0.21). A 10 mm average daily increase in rainfall in the six months prior to the survey (NOAA4) was non-significantly associated with a 3% increase in prevalence. Increases of 100 m in altitude were associated with a non-significant 4% decrease in prevalence.

All LLIN variables serve as a proxy for appropriate LLIN usage, and were protective, but not significant in univariate analysis (p=0.10-0.24). LLIN associations were more protective than all net association, which included LLINs and non-treated nets.

Table 1: Prevalence, LLIN Coverage, and Remote Sensed Variable Means by Survey Period in Amhara, Oromiya, and SNNP (2006 Baseline Survey)

Variable	Period 1*^ Means (N = 128)	Period 2*^ Means (N = 95)	Mean for all kebels (N = 223)		
Malaria Prevalence^^	5.0%	2.4%	4.2%	**Variable description**	
LLIN coverage**	16.1%	28.7%	21.5%	**(% HH with at least one LLIN)**	
NDVI from 16 day MODIS composites at 250 meter resolution				**Avg. Normalized Difference Vegetation Index.**	**Ecological/Biological Justification**
NDVI1**	0.43	0.45	0.44	0-16 days prior to survey	Humidity
NDVI2	0.45	0.46	0.45	16-32 days prior to survey	Humidity, Larval sites
NDVI3	0.44	0.45	0.45	0-32 days prior to survey	Humidity, Larval sites
NDVI4	0.47	0.47	0.47	0-48 days prior to survey	Humidity, Larval sites
Night time LST from 8 day MODIS composites at 1 km resolution in degrees K				**Average night time Land Surface Temperature (K).**	
LST1	285.3	285.1	285.2	0-16 days prior to survey	Parasite development, Mosquito survival
LST2	285.5	285.8	285.6	16-32 days prior to survey	Larval development
LST3*,**	285.4	285.5	285.4	0-32 days prior to survey	Larval development, Parasite development, Mos. survival
LST4*	285.7	285.3	285.6	0-48 days prior to survey	Larval development, Parasite development, Mos. survival
African Rainfall Climatology from the NOAA Famine Early Warning System daily estimates in mm.				**Avg daily rainfall from the National Oceans and Atmospheric Admin.**	
NOAA1	7.6	9.2	8.1	Month prior to survey	Humidity, Larval sites
NOAA2	17.7	13.7	16.4	2 Months prior to survey	Humidity, Larval sites
NOAA3^^	22.52	16.82	20.68	3 months prior to survey	Humidity, Transmission during peak season, Larval sites
NOAA4,^^**	52.60	32.84	46.24	6 months prior to survey	Transmission during seasonal rains and during peak transmission season
NOAA5^^	42.11	33.64	39.38	9 months prior to survey	Annual suitability

*** Significant (p < 0.5) in univariate analysis**

**** Remained in final model**

***^ Kebeles surveyed between 12 Dec and 31 Dec 2006 were grouped into Period 1, and were assigned a survey date of 22 Dec 2006 for comparison with remote sensed variables, while those surveyed between 1 Jan and 7 Feb 2007 were assigned a survey data of 18 Jan 2007.**

^^ Significantly different between the two Periods (p<0.05)

As NDVI and NDWI were highly correlated, only NDVI values are presented.

Table 2: Univariate Associations in Poisson Regression Comparing Prevalence with Village Level Demographic, Behavioural, Malaria Control, and Environmental Factors in Amhara, Oromiya, and SNNP (2006 Baseline Survey)

Parameter	OR **	CL	P-value*
Survey Period* (Period 1 is referent)	0.77	0.65-0.92	0.003
Sqrt of Population Density (per person per km²)	1.02	0.98-1.06	0.32
Population Density* (per person per km²)	1.001	1.000-1.002	0.03
Average Asset Index per kebele	0.94	0.75-1.18	0.57
Average individual Net use last night (per individual)	0.74	0.45-1.21	0.23
Average individual LLIN use last night (per individual)	0.61	0.32-1.14	0.12
Ratio of Nets to HH members (per net)	0.74	0.33-1.63	0.45
Ratio of LLINs to HH members (per net	0.53	0.18-1.52	0.24
Number of Nets per HH (per net)	0.84	0.67-1.06	0.14
Number of LLINs per HH (per net)	0.79	0.59-1.06	0.11
% of HHs with at least 1 Net (per 20% increase)	0.82	0.55-1.21	0.31
% of HHs with at least 1 LLIN (per 20% increase)	0.67	0.41-1.09	0.10
IRS 12 Months	0.72	0.41-1.28	0.27
IRS 6 Months	1.20	0.58-2.47	0.62
Hydoline within 3000m	1.45	0.86-2.45	0.16
Kebele Altitude (per meter)	1.000	0.999-1.000	0.42
LST1 (per K)	1.07	1.00-1.16	0.054
LST2 (per K)	1.06	0.99-1.14	0.12
LST3 (per K)*	1.08	1.00-1.16	0.049
LST4 (per K)*	1.09	1.01-1.17	0.03
NDVI1	0.44	0.12-1.58	0.21
NDVI2	0.64	0.19-2.15	0.47
NDVI3	0.53	0.15-1.86	0.32
NDVI4	0.62	0.18-2.13	0.45
NOAA1 (per mm)	1.00	0.98-1.01	0.59
NOAA2 (per mm)	1.00	0.99-1.01	0.56
NOAA3 (per mm)	1.00	0.99-1.01	0.46
NOAA4 (per mm)	1.01	1.00-1.02	0.13
NOAA5 (per mm)	1.00	0.99-1.02	0.59

Multivariate Associations in Poisson Regression

LST3 was included as the base environmental model from discriminant analysis. Population density was included and retained. In multivariate analysis, a positive relationship between NOAA4 and prevalence became significant when controlling for population density, with a 10 mm increase in average daily rainfall corresponding to a 17% increase in risk.

A negative relationship between prevalence and NDVI1 became significant when controlling for NOAA4. NDVI1 correlates strongly with rainfall in the month proceeding the survey, which was also negatively correlated with prevalence.

The average number of LLINs owned per household, the average number of households owning at least one LLIN (LLIN coverage), and the proportion of the household who used a LLIN the night prior to the survey became significant predictors when controlling for temperature and population density. With the inclusion of rainfall, vegetation, and IRS, LLIN coverage remained significantly protective (PR 0.6 per 20% increase coverage, p=0.04).

Altitude correlates strongly with temperature. If included in place of temperature in our final model, altitude would have been significantly protective (p = 0.003).

Despite not meeting the selection criteria, we included indoor residual spraying (IRS) within the last 12 months for comparison with previous analysis. When stratified on altitude instead of temperature, IRS is borderline significant (p=0.056).

Table 3: Fully Adjusted Associations in Poison Regression Between Malaria Prevalence, Demographics, Net Coverage, and Environmental Factors in Amhara, Oromiya, and SNNP (2006 Baseline Survey)

Parameter (In the order selected in stepwise regression)	Prevalence Ratio	CL	P-value
LST3**	1.13	1.05 -1.22	0.002
Population density**	1.002	1.001 -1.003	0.0004
NOAA4 (per mm)**	1.02	1.01 -1.03	0.005
NDVI1**	0.15	0.04 -0.60	0.008
Proportion of HH with at least 1 LLIN (per 20% increase)*	0.60	0.37 -0.98	0.041
IRS within last 12 Months	0.67	0.39 -1.18	0.166

* Significant (p<0.05)
** Significant (p<0.01)

Discussion

Although aggregation by region indicates that Amhara and SNNP have higher prevalence than Oromiya, we saw no evidence of global clustering when considering all three regions together. The clustering on the eastern border of SNNP was not detected by either Moran's I and LISA

This lack of pattern changed when we moved down to a finer scale, where significant clustering in Oromiya and SNNP is driven by the large cluster on the eastern border of SNNP. This cluster is found in an area of moderate temperature, high population density, high bednet use the night before the survey, and high cloud cover in the month leading up to the survey period. The high net usage in the night prior to the survey is surprising, and demonstrates the challenge in identifying the association between appropriate bed net use [77] and risk of malaria through a cross-sectional study—increased mosquito population due to suitable habitat in the months leading up to the survey and the corresponding nuisance biting together with changing perceptions of risk could result in both increased prevalence and increased net use. Investigation into mosquito biting patterns, net quality, net hanging techniques, and consistent LLIN use in other areas of high prevalence may provide further insight into the effectiveness of LLIN distribution strategies.

As this cluster is found in one of the most densely populated areas in our study area, we suggest potential micro-environment factors which may be unique to this area—such as peri-urban agriculture––may be important to local transmission. Further, we suggest investigation into the role existing immunity together with population movement may play in malaria transmission in more densely populated regions. Finally, it may be that a certain critical population density in a rural environment facilitates transmission, as mosquitoes have larger host populations on which to feed, and less distance to travel for next feeding.

Investigation of risk factors at the individual level [62] suggested a significant association (positive or negative) between the likelihood of a person being infected with malaria and: altitude, the asset index, rainfall, and the number of LLINs per house. Previous research on ecological factors associated with malaria transmission has included altitude {Omumbo, 2005 #178;Graves, 2009 #63;Bodker, 2003 #170;Kulkarni, 2010 #173}, urban or rural stratification [68], landcover [75,78,79], night time temperature [57,69,80], precipitation [57,59,69,70,81,82], vegetation [59,69,78], proximity to water [83], population density [24,84,85], and clinic locations [86], and hydrology [24,87].

By aggregating to the village level, we have attempted to tease out some of the relative importance of malaria control factors and environmental factors in driving the spatial heterogeneity of malaria in Ethiopia.

LLIN coverage or use in our study are protective when controlling for environmental parameters. This findings add to existing evidence [63,88,89,90] on the importance of LLINs for malaria control.

Regarding the environmental factors, our findings suggest that LST provides a better indication of malaria risk than altitude, which has historically been used to prioritize net coverage and treatment distributions. Further research and increasing public access to high quality environmental data may support classification of risk based on LST rather than altitude. Altitude serves as a proxy for temperature, which affects both parasite and vector biology. More research is needed into how annual variation in temperature may shift the regions at risk, which could increase or decrease the altitude threshold for programmatic implementation for a specific year.

Higher rainfall during the peak rains corresponded with increased risk, though rainfall in the month prior to the survey corresponded with decreased risk. This may reflect heavy rains washing out larval development sites. Finer temporal scale investigation of daily rainfall patterns, specifically maximum daily rainfall, is needed. These findings highlight the complexity of the relationship between rainfall and mosquito habitat, and the importance that sustained transmission in the weeks leading up to a survey can have on prevalence estimates.

There are several possible interpretations for the negative association between NDVI and prevalence. Nearby water tend to lower the average NDVI around a village. Finally, NDVI may serve as a unreliable predicator, as it reflects an aggregate of vegetation, water bodies, humidity, rainfall (through vegetation growth), and land use, which may be difficult to disentangle. The tasseled cap transformations discussed in our further research section should be more informative.

The differences in regional prevalence are unlikely to be due to survey timing––the survey's time frame was short, and the average LST (the strongest predictor) is similar over the two periods.

The large scale up in net distribution since the 2006 Baseline survey has resulted in significant increases in net ownership. Further research into changes or a lack of change in these environmental parameters may provide stronger evidence supporting the claim that the net distribution campaigns have caused the declining prevalence of malaria nationwide.

Public Health Implications

The 2006 Baseline survey covers the three largest regions containing 80% of the country's population [91]. As such, it provides an estimate of prevalence which can be extrapolated nationally. Our analysis provides a context for evaluating the impact increased net coverage has played on decreasing malaria prevalence, while controlling for environment.

As Ethiopia continues to roll out LLINs for malaria prevention, an updated malaria risk map that considers demographics, behaviour, and knowledge about malaria in addition to weather and ecology is needed to target further, focused interventions [92]. This cross-sectional survey provides an important piece of the risk stratification through these three regions. A more complete risk map would require quality longitudinal data on national cases. Longitudinal data tends, however, to be subject to biases from reporting, access to care, and quality of treatment, while a survey is not.

As the Ethiopian FMOH sees further progress toward the Roll Back Malaria goals, we expect malaria transmission to grow more focal. After validation, the association between LST, Rainfall, and NDVI and malaria prevalence at fine scales may help drive sentinel surveillance or prioritization of net distribution or redistribution.

Lack of clear clustering in Amhara and Oromiya demonstrates the heterogeneity of malaria within Ethiopia. Given the clustering of we saw in SNNP, we emphasize the need examine surveillance data and respond to outbreaks at the zonal and woreda level, as region-wide prevalence estimates may fail to identify local areas of high transmission.

Currently the FMOH uses altitude to define areas of high and low priority. Programmatically, altitude as a proxy for temperature provides a beginning for stratifying malaria risk. Our results indicate that although altitude serves as a proxy for temperature, temperature itself is the driving factor, and should be included in programmatic considerations—especially for those areas whose classifications are borderline more-malarious or less-malarious based on altitude. Better intervention targeting should allow health resources to be more efficiently used. Although temperature (like many other environmental parameters) measured through remote sensed satellite imagery can be problematic due to contamination from cloud cover, recent collaboration by Ethiopia's National Meteorological Agency, the IRI, and the Tropical Applications of Meteorology using Satellite Data (TAMSAT) toward a comprehensive climate database holds great promise [93].

However, even with the best environmental data available, risk maps will be imprecise [69]. As Ethiopia continues to demonstrate progress in malaria control and damage to existing nets accumulates, perceptions of benefit from net use maydecrease [77]. Existing malaria control, net ownership, quality and use, a region's ability to respond rapidly to outbreaks, and its health systems strength should be considered to prioritize net re-distribution and education in addition to habitat suitability.

As we conclude, it is important to remember that all models are, in some way, flawed. Yet, to the extent that they meet necessary assumptions and provide better understanding of disease distribution and dynamics, they can be helpful in identifying risk factors and trends, and in driving efficient use of resources for disease control.

Limitations

There is heterogeneity of peak transmission seasons between different years and regions. Cross-sectional data, even timed to be during the peak malaria transmission, cannot capture this temporal variation.

To properly evaluate the causal link between LLIN use and malaria, we would need to measure appropriate LLIN use in the week to two weeks (based on parasite incubation) prior to evaluating infection. We proxy this measure by asking about current use and current coverage, and assuming that current use is reflective of prior use. Asking about net use in the previous month would be subject to recall bias and would thus be less accurate than the existing questions of net ownership and use in the night before the survey.

The relationship between malaria and environment is complex and not fully understood. We have attempted to proxy certain environmental components which play key roles in malaria transmission through remote sensed data. There is potential for error first in our choice of proxies––precipitation and rainfall do not completely correlate with humidity and breeding sites [94]. Proximity to lakes, rivers, or drainage does not necessarily indicate suitable breeding sites [83]. Habitats created through wells, stored water, irrigation, micro-dams, or puddles from pits and tire tracks play large roles in mosquito abundance and corresponding transmisison, and were not detectable through our remote sensed covariates [92].

Further Research

Further research is needed to overcome our limitations regarding habitat measurement. Other remote sensed indexes should be considered, such as the tasseled cap transformations which measure landscape brightness and greenness, and soil moisture. Aspect, slope, soil moisture, and soil texture may play important roles in formation of larval habitat. Further, non-exponential, linear, and non-linear relationships between remote sensed parameters and prevalence should be considered.

Due to over-dispersion, zero-inflated poisson or negative binomial regression should be considered in further analysis. Dichotomous logistic regression models based on presence or absence of any malaria, and based on low (<3%) and high (>3%) were considered, and will be included in further analysis.

Potential interactions which may further elucidate patterns of transmission include LLIN use and temperature; or rainfall and temperature.

Residuals from our final model should be mapped and examined for clusters of poor fit, which would indicate spatial autocorrelation.

By aggregating to the village level, village estimates are treated as measured values rather than as estimates with variability. This limitation would lead to narrower confidence intervals than appropriate. Hierarchical modeling accounting for survey design would allow us to incorporate the variability at the individual and household level with the village level environmental factors of interest, but brings with it its own difficulties in interpretation [95,96].

Finally, validation of our results together with more data from additional times, together with data from various ranges of malaria control implementation, are necessary to support using environmental risk factors such as temperature in addition to elevation to target future interventions.

Maps

Fig1: Malaria Prevalence and Elevation

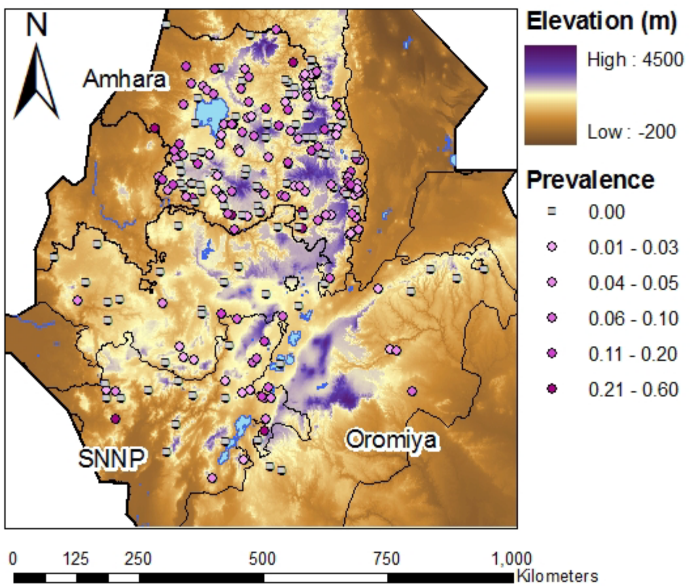


Fig 1b: Anthropogenic Biomes

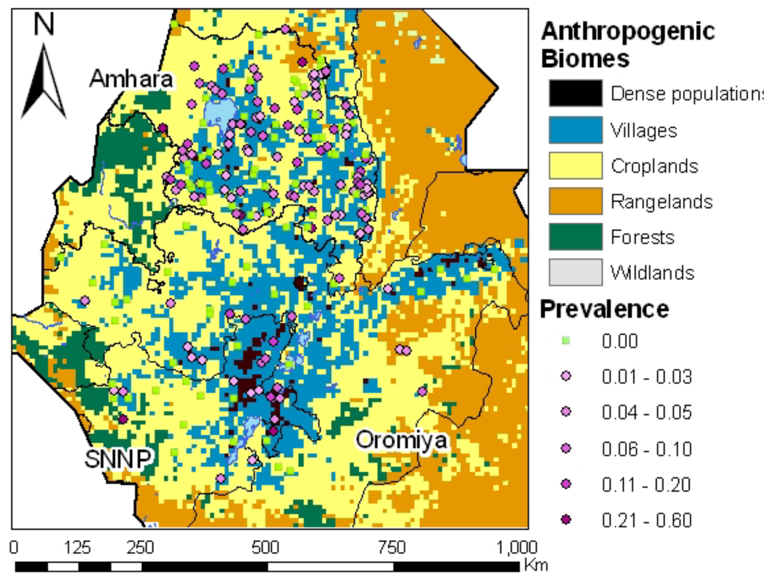


Fig2a: Daily Rainfall (NOAA4) in Period 1

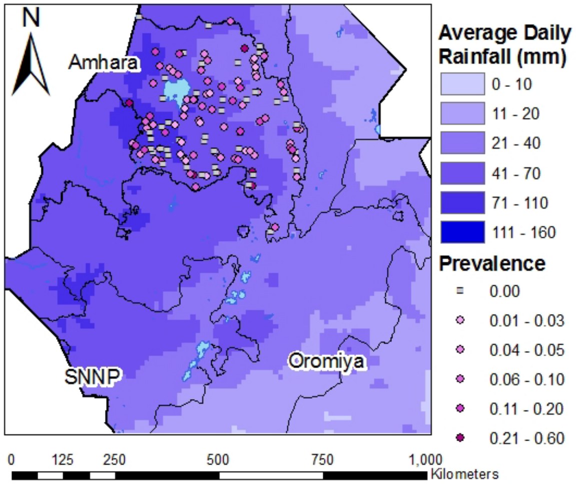


Fig2b: Daily Rainfall (NOAA4) in Period 2

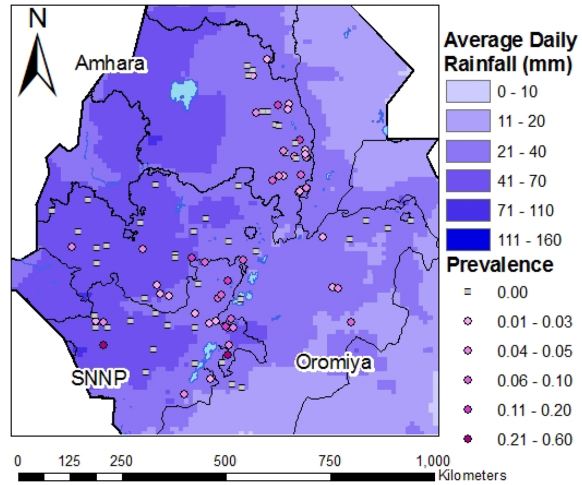


Fig3a: Average Vegetation (NDVI1) in Period 1

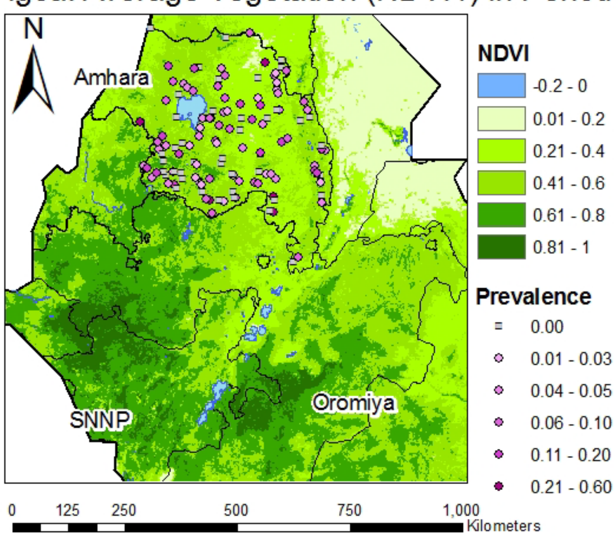


Fig3b: Average Vegetation (NDVI1) in Period 2

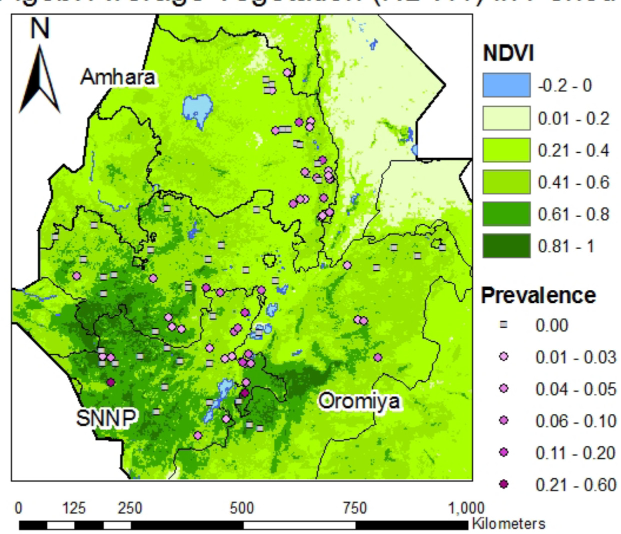


Fig4a: Temperature in C (LST3) in Period 1

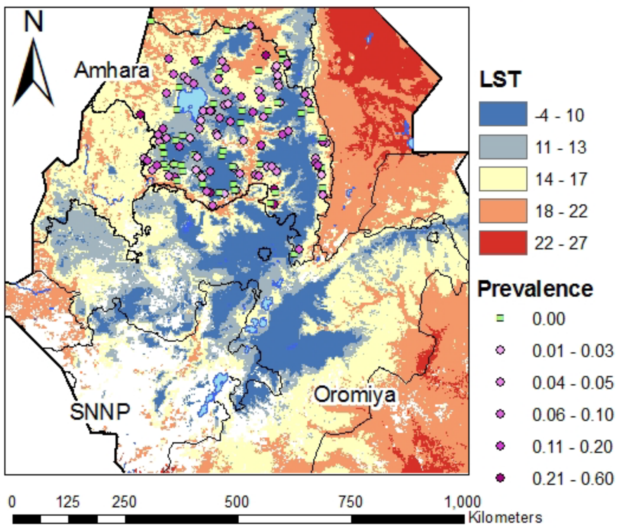


Fig4b: Temperature in C (LST3) in Period 2

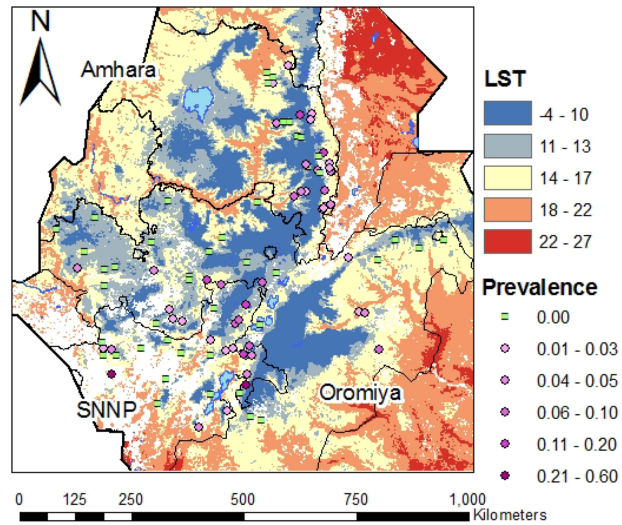


Fig 5: Malaria Prev and Population Density

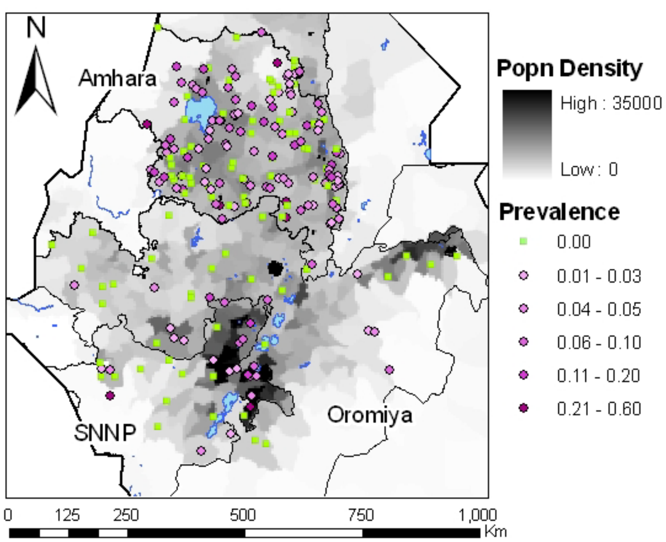


Fig 6a: Presence or Absence of Malaria per Kebele

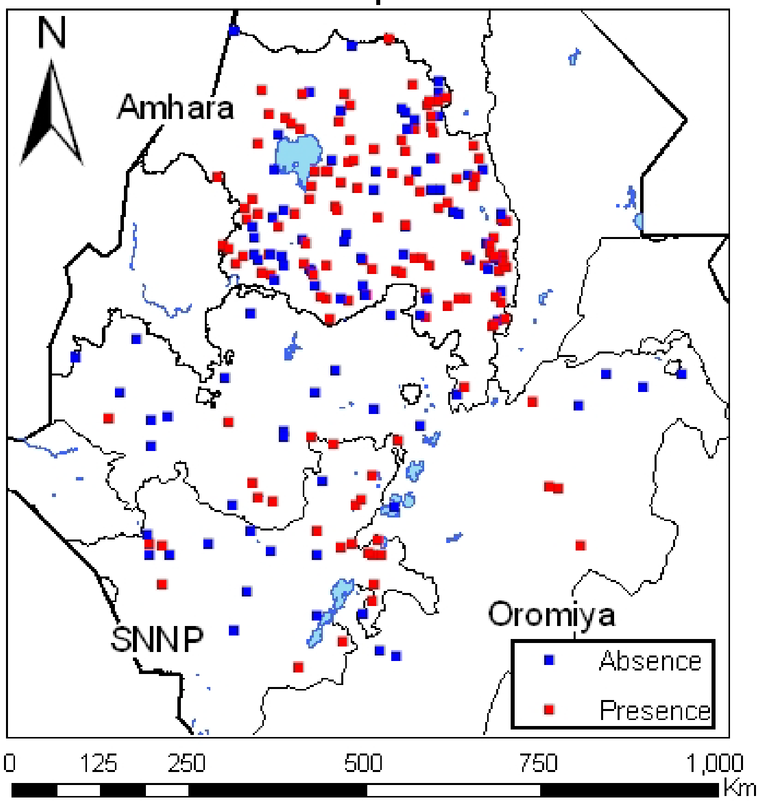


Fig 6b: Prevalence of Malaria per Kebele

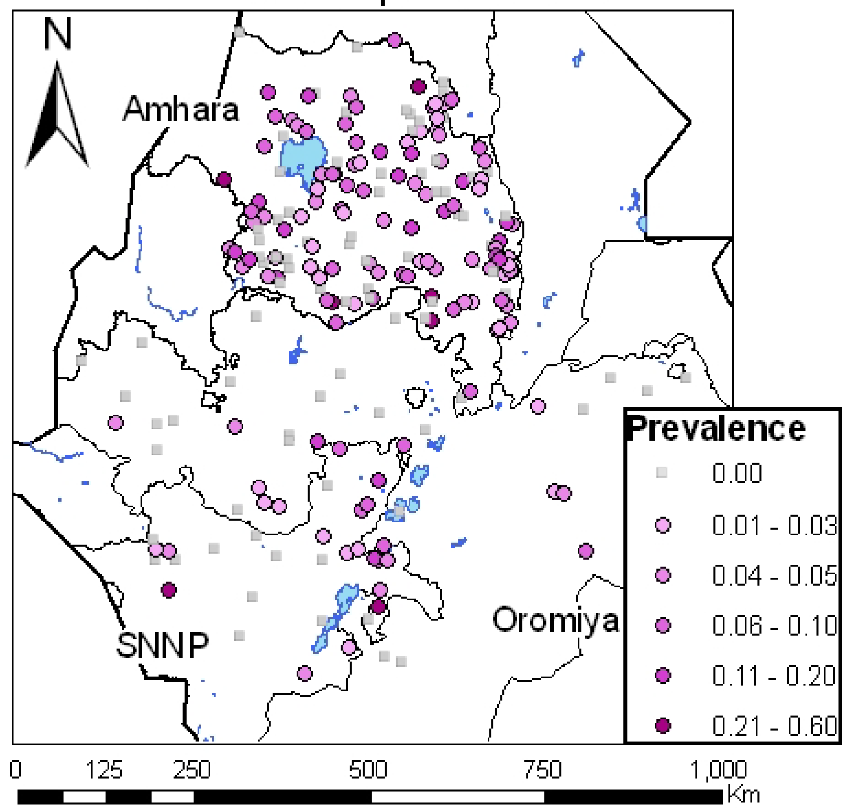
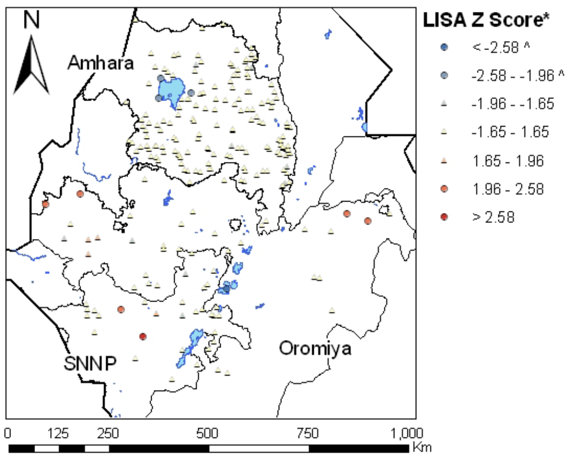
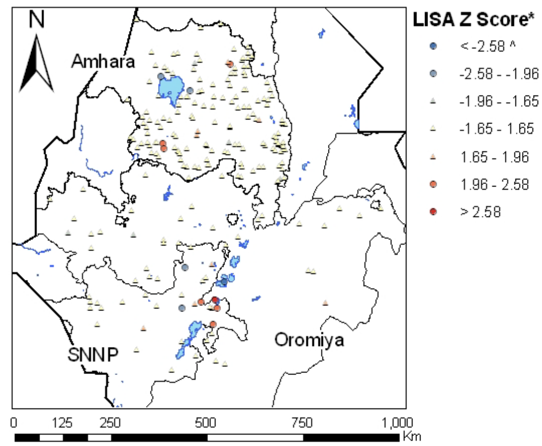


Fig 7a: LISA for Presence/Absence of Malaria Across All Regions



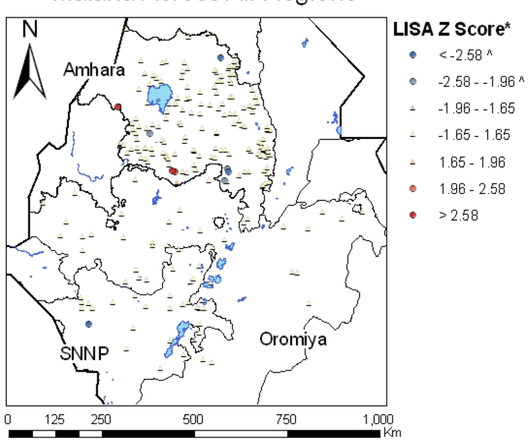
* Negative values indicate areas of low-low concentrations.
Positive values indicate clustering of hi-hi values.
^ Significant at $p < 0.05$.

Fig 7b: LISA for Presence/Absence of Malaria Amhara alone and Oromiya with SNNP



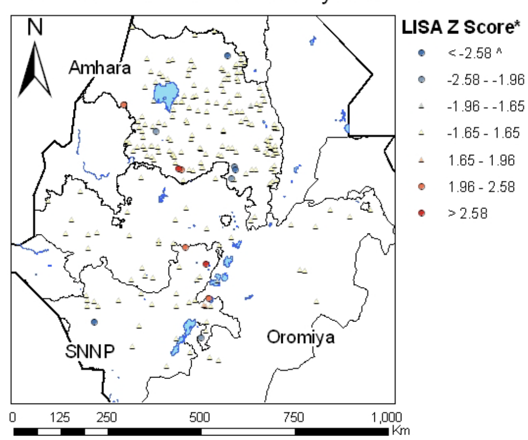
* Negative values indicate areas of low-low concentrations.
Positive values indicate clustering of hi-hi values.
^ Significant at $p < 0.05$.

Fig 7d: LISA for Prevalence of Malaria Across All Regions



* Negative values indicate areas of low-low concentrations.
Positive values indicate clustering of hi-hi values.
^ Significant at $p < 0.05$.

Fig 7d: LISA for Prevalence of Malaria Amhara alone and Oromiya with SNNP



* Negative values indicate areas of low-low concentrations.
Positive values indicate clustering of hi-hi values.
^ Significant at $p < 0.05$.

Funding

The 2006 Baseline survey was funded by The Carter Center. Survey staff were provided by the Amhara, Oromia and SNNP regional health bureaux [62].

Author details

Marc Cunningham (MC) is completing his Master of Public Health in the Epidemiology department at Rollins School of Public Health. Uriel Kitron (UK), in addition to being Professor and Chair of the Department of Environmental Studies at Emory University, is a Professor of Emory's Program in Behavior, Ecology and Evolution (PBEE), of Environmental Health, and of Epidemiology at Rollins School of Public Health. Patricia Graves (PG) is an Program Epidemiologist for The Carter Center Malaria Control Program, and has 25 years of experience in applied research in vector-borne disease epidemiology.

Authors' contributions

PG and MC identified the research question. MC drafted and edited the manuscript, and performed all analysis. PG provided expert knowledge and survey background. UK provided technical assistance in spatial analysis techniques and vector biology and ecology. UK and PG edited the manuscript, and will provide additional feedback prior to publication. All opinions in this manuscript reflect those held MC.

Competing interests

The Authors have no competing interests.

Bibliography

1. Tobler WR (1970) Computer Movie Simulating Urban Growth in Detroit Region. Economic Geography 46: 234-240.

2. Sui DZ (2004) Tobler's first law of geography: A big idea for a small world? Annals of the Association of American Geographers 94: 269-277.

3. Kerr JT, Ostrovsky M (2003) From space to species: ecological applications for remote sensing. Trends in Ecology & Evolution 18: 299-305.

4. Robinson TP (2000) Spatial statistics and geographical information systems in epidemiology and public health. Advances in Parasitology, Vol 47 47: 81-128.

5. Moran PAP (1950) Notes on Continuous Stochastic Phenomena. Biometrika 37: 17-23.

6. Cliff AD, Ord JK (1981) Spatial Processes: Models and Applications. London: Pion.

7. Kitron U, Kazmierczak JJ (1997) Spatial analysis of the distribution of Lyme disease in Wisconsin. American Journal of Epidemiology 145: 558-566.

8. Ripley BD (1976) The second-order analysis of secondary point processes. Journal of Applied Probability 13: 255-266.

9. Ripley BD (1981) Spatial statistics. New York: Wiley. x, 252 p. p.

10. Anselin L (1995) Local Indicators of Spatial Association - Lisa. Geographical Analysis 27: 93-115.

11. Getis A, Ord JK (1992) The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis 24: 189-206.

12. Getis A, Ord JK (1996) Local spatial statistics: An overview. In: Longley P, Batty M, editors. Spatial Analysis: Modelling in a GIS Environment. Cambridge: Geoinformation International. pp. 261-277.

13. Holveck JC, Ehrenberg JP, Ault SK, Rojas R, Vasquez J, et al. (2007) Prevention, control, and elimination of neglected diseases in the Americas: Pathways to integrated, inter-programmatic, inter-sectoral action for health and development. Bmc Public Health 7: -.

14. Pampana E (1969) A textbook of malaria eradicaition. London: Oxford University Press.

15. Hay SI (2000) An overview of remote sensing and geodesy for epidemiology and public health application. Advances in Parasitology, Vol 47 47: 1-35.

16. Elliott P, Wartenberg D (2004) Spatial epidemiology: Current approaches and future challenges. Environmental Health Perspectives 112: 998-1006.

17. Huete A, Didan K, Miura T, Rodriguez EP, Gao X, et al. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83: 195-213.

18. Gao BC (1996) NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58: 257-266.

19. Hay SI, Lennon JJ (1999) Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. Tropical Medicine & International Health 4: 58-71.

20. Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled seamless SRTM data V4: International Centre for Tropical Agriculture (CIAT).

21. Kitron U (2000) Risk maps: Transmission and burden of vector borne diseases. Parasitology Today 16: 324-325.

22. Wood BL, Beck LR, Lobitz BM, Bobo MR (2000) Education, outreach and the future of remote sensing in human health. Advances in Parasitology, Vol 47 47: 331-344.

23. Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW (2004) The global distribution and population at risk of malaria: past, present, and future. Lancet Infectious Diseases 4: 327-336.

24. Zhou G, Munga S, Minakawa N, Githeko AK, Yan G (2007) Spatial relationship between adult malaria vector abundance and environmental factors in western Kenya highlands. American Journal of Tropical Medicine and Hygiene 77: 29-35.

25. James FC, Mcculloch CE (1990) Multivariate-Analysis in Ecology and Systematics - Panacea or Pandora Box. Annual Review of Ecology and Systematics 21: 129-166.

26. Hay SI, Snow RW, Rogers DJ (1998) From predicting mosquito habitat to malaria seasons using remotely sensed data: Practice, problems and perspectives. Parasitology Today 14: 306-313.

27. Baylis M, Rawlings P (1998) Modelling the distribution and abundance of Culicoides imicola in Morocco and Iberia using climatic data and satellite imagery. Archives of Virology: 137-153.

28. Anselin L, Rey S (1991) Properties of Tests for Spatial Dependence in Linear-Regression Models. Geographical Analysis 23: 112-131.

29. Diggle P, Moyeed R, Rowlingson B, Thomson M (2002) Childhood malaria in the Gambia: a case-study in model-based geostatistics. Journal of the Royal Statistical Society Series C-Applied Statistics 51: 493-506.

30. Duchateau L, Kruska RL, Perry BD (1997) Reducing a spatial database to its effective dimensionality for logistic-regression analysis of incidence of livestock disease. Preventive Veterinary Medicine 32: 207-218.

31. Best NG, Ickstadt K, Wolpert RL (2000) Spatial Poisson regression for health and exposure data measured at disparate resolutions. Journal of the American Statistical Association 95: 1076-1088.

32. Jarup L, Best N, Toledano MB, Wakefield J, Elliott P (2002) Geographical epidemiology of prostate cancer in Great Britain. International Journal of Cancer 97: 695-699.

33. Lewsey JD, Thomson WM (2004) The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. Community Dent Oral Epidemiol 32: 183-189.

34. Getis A (1990) Screening for Spatial Dependence in Regression-Analysis. Papers of the Regional Science Association 69: 69-81.

35. Haining R (1990) Spatial Data Analysis in the Social and Envirnmental Sciences. Cambridge: Cambridge University Press.

36. Cressie NAC (1993) Statistics for Spatial Data. New York: Wiley and Sons.

37. Graham AJ, Atkinson PM, Danson FM (2004) Spatial analysis for epidemiology. Acta Tropica 91: 219-225.

38. Fotheringham AS, Rogerson PA (1993) Gis and Spatial Analytical Problems. International Journal of Geographical Information Systems 7: 3-19.

39. Moellering H, Tobler W (1972) Geographical Variances. Geographical Analysis 4: 34-50.

40. Haining RP, Kerry R, Oliver MA (2010) Geography, Spatial Data Analysis, and Geostatistics: An Overview. Geographical Analysis 42: 7-31.

41. Waller LA, Gotway CA (2004) Applied Spatial Statistics for Public Health. New Jersey: John Wiley & Sons.

42. Kitron U (1998) Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. J Med Entomol 35: 435-445.

43. Cliff AD, Hagget P (1988) Atlas fo the Distribution of Diseases: Analytical Approaches to Epidemiological Data Oxford: Blackwell.

44. Ostfeld RS, Glass GE, Keesing F (2005) Spatial epidemiology: an emerging (or re-emerging) discipline. Trends in Ecology & Evolution 20: 328-336.

45. Levins R (1966) Strategy of Model Building in Population Biology. American Scientist 54: 421-&.

46. WHO (2010) World Malaria Report. 89-91 p.

47. PMI (2011) Malaria Operations Plan: Ethiopia FY2011.

48. MIS (2007) Ethiopia Malaria Indicator Survey 2007.

49. FMOH (2006) National Five Year Malaria Control Plan

50. Ribeiro JMC, Seulu F, Abose T, Kidane G, Teklehaimanot A (1996) Temporal and spatial distribution of anopheline mosquitos in an Ethiopian village: Implications for malaria control strategies. Bulletin of the World Health Organization 74: 299-305.

51. Hay SI, Omumbo JA, Craig MH, Snow RW (2000) Earth observation, geographic information systems and Plasmodium falciparum malaria in sub-Saharan Africa. Advances in Parasitology, Vol 47 47: 173-215.

52. Service MW (1971) Studies on Sampling Larval Populations of Anopheles-Gambiae Complex. Bulletin of the World Health Organization 45: 169-&.

53. Jima D, Hwang J, Getachew A, Bilak H, Shargie EB, et al. (2008) Achievements in Malaria Control in Ethiopia: Results from Malaria Indicator Survey, 2007. American Journal of Tropical Medicine and Hygiene 79: 64-64.

54. BP N (1935) The influence of temperature on the development of malaria plasmodia in the mosquito. Tr Pasteur Inst Epidem Bakt 2: 108-109.

55. Macdonald G (1956) Epidemiological basis of malaria control. Bulletin of the World Health Organization 15: 613–626.

56. Dogan HM, Cetin I, Egri M (2010) Spatiotemporal change and ecological modelling of malaria in Turkey by means of geographic information systems. Transactions of the Royal Society of Tropical Medicine and Hygiene 104: 726-732.

57. Paaijmans KP, Jacobs AFG, Takken W, Heusinkveld BG, Githeko AK, et al. (2008) Observations and model estimates of diurnal water temperature dynamics in mosquito breeding sites in western Kenya. Hydrological Processes 22: 4789-4801.

58. Rogers DJ, Randolph SE, Snow RW, Hay SI (2002) Satellite imagery in the study and forecast of malaria. Nature 415: 710-715.

59. Thomson MC, Connor SJ, Milligan P, Flasse SP (1997) Mapping malaria risk in Africa: What can satellite data contribute? Parasitology Today 13: 313-318.

60. Shargie EB, Ngondi J, Graves PM, Getachew A, Hwang J, et al. (2010) Rapid increase in ownership and use of long-lasting insecticidal nets and decrease in prevalence of malaria in three regional States of ethiopia (2006-2007). J Trop Med 2010.

61. Emerson PM, Ngondi J, Biru E, Graves PM, Ejigsemahu Y, et al. (2008) Integrating an NTD with One of "The Big Three'': Combined Malaria and Trachoma Survey in Amhara Region of Ethiopia. Plos Neglected Tropical Diseases 2: -.

62. Graves PM, Richards FO, Ngondi J, Emerson PM, Shargie EB, et al. (2009) Individual, household and environmental risk factors for malaria infection in Amhara, Oromia and SNNP regions of Ethiopia. Transactions of the Royal Society of Tropical Medicine and Hygiene 103: 1211-1220.

63. Shargie EB, Gebre T, Ngondi J, Graves PM, Mosher AW, et al. (2008) Malaria prevalence and mosquito net coverage in Oromia and SNNPR regions of Ethiopia. Bmc Public Health 8.

64. Jima D, Getachew A, Bilak H, Steketee RW, Emerson PM, et al. (2010) Malaria indicator survey 2007, Ethiopia: coverage and use of major malaria prevention and control interventions. Malaria Journal 9: -.

65. Hwang J, Graves PM, Jima D, Reithinger R, Kachur SP, et al. (2010) Knowledge of Malaria and Its Association with Malaria-Related Behaviors-Results from the Malaria Indicator Survey, Ethiopia, 2007. PLoS One 5.

66. Bodker R, Akida J, Shayo D, Kisinza W, Msangeni HA, et al. (2003) Relationship between altitude and intensity of malaria transmission in the Usambara Mountains, Tanzania. Journal of Medical Entomology 40: 706-717.

67. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, et al. (2009) A World Malaria Map: Plasmodium falciparum Endemicity in 2007. Plos Medicine 6: -.

68. Hay SI, Guerra CA, Tatem AJ, Atkinson PM, Snow RW (2005) Urbanization, malaria transmission and disease burden in Africa. Nature Reviews Microbiology 3: 81-90.

69. Omumbo JA, Hay SI, Snow RW, Tatem AJ, Rogers DJ (2005) Modelling malaria risk in East Africa at high-spatial resolution. Tropical Medicine & International Health 10: 557-566.

70. Thomson MC, Mason SJ, Phindela T, Connor SJ (2005) Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. American Journal of Tropical Medicine and Hygiene 73: 214-221.

71. Ashton RA, Kefyalew T, Tesfaye G, Pullan RL, Yadeta D, et al. (2011) School-based surveys of malaria in Oromia Regional State, Ethiopia: a rapid survey method for malaria in low transmission settings. Malar J 10: 25.

72. Tilaye T, Deressa W (2007) Prevalence of urban malaria and assocated factors in Gondar Town, Northwest Ethiopia. Ethiop Med J 45: 151-158.

73. Deichmann U, Balk D, Yetman G (2001) Transforming Population Data for Interdisciplinary Usages: From census to grid.

74. Love TB, V. Kumar, P. Xie, Thiaw W (2004) A 20-Year Daily Africa Precipitation Climatology Using Satellite and Gauge Data. AMS.

75. Ellis EC, Ramankutty N (2008) Putting people in the map: anthropogenic biomes of the world. Frontiers in Ecology and the Environment 6: 439-447.

76. Pfeiffer D, Robinson T, Stevenson M, Stevens K, Rogers DJ, et al. (2008) Spatial Analysis in Epidemiology. New York: Oxford University Press, Inc.

77. Ngondi J, Graves P, Gebre T, Mosher A, Shargie EB, et al. (2011) Which nets are being used: factors associated with mosquito net use in Amhara, Oromia and Southern Nations, Nationalities and Peoples' Regions of Ethiopia. Malaria Journal 10.

78. Kulkarni MA, Desrochers RE, Kerr JT (2010) High Resolution Niche Models of Malaria Vectors in Northern Tanzania: A New Capacity to Predict Malaria Risk? PLoS One 5: -.

79. Bogh C, Lindsay SW, Clarke SE, Dean A, Jawara M, et al. (2007) High spatial resolution mapping of malaria transmission risk in The Gambia, West Africa, using landsat TM satellite imagery. American Journal of Tropical Medicine and Hygiene 76: 875-881.

80. Minakawa N, Sonye G, Mogi M, Yan G (2004) Habitat characteristics of Anopheles gambiae s.s. larvae in a Kenyan highland. Medical and Veterinary Entomology 18: 301-305.

81. Lindsay SW, Bayoh MN (2004) Mapping members of the Anopheles gambiae complex using climate data. Physiological Entomology 29: 204-209.

82. Huffman G (2007) The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. J Hydrometeor 8: 38-55.

83. Ageep TB, Cox J, Hassan MM, Knols BGJ, Benedict MQ, et al. (2009) Spatial and temporal distribution of the malaria mosquito Anopheles arabiensis in northern Sudan: influence of environmental factors and implications for vector control. Malaria Journal 8: -.

84. Tatem AJ, Guerra CA, Kabaria CW, Noor AM, Hay SI (2008) Human population, urban settlement patterns and their impact on Plasmodium falciparum malaria endemicity. Malaria Journal 7: -.

85. Hay SI, Noor AM, Nelson A, Tatem AJ (2005) The accuracy of human population maps for public health application. Tropical Medicine & International Health 10: 1073-1086.

86. Martin C, Curtis B, Fraser C, Sharp B (2002) The use of a GIS-based malaria information system for malaria research and control in South Africa. Health & Place 8: 227-236.

87. Bomblies A, Duchemin JB, Eltahir EAB (2008) Hydrology of malaria: Model development and application to a Sahelian village. Water Resources Research 44: -.

88. Charlwood JD, Alcantara J, Pinto J, Sousa CA, Rompao H, et al. (2005) Do bednets reduce malaria transmission by exophagic mosquitoes? Transactions of the Royal Society of Tropical Medicine and Hygiene 99: 901-904.

89. Deribew A, Alemseged F, Birhanu Z, Sena L, Tegegn A, et al. (2010) Effect of training on the use of long-lasting insecticide-treated bed nets on the burden of malaria among vulnerable groups, south-west Ethiopia: baseline results of a cluster randomized trial. Malar J 9: 121.

90. Ghebreyesus TA (2010) Achieving the health MDGs: country ownership in four steps. Lancet 376: 1127-1128.

91. Health Information Processing and Documentation Team FE (2005) Health and Health Related Indicators, 2004/2005. Addis Ababa.

92. Carter R, Mendis KN, Roberts D (2000) Spatial targeting of interventions against malaria. Bulletin of the World Health Organization 78: 1401-1411.

93. Thomson MC, Connor SJ, Zebiak SE, Jancloes M, Mihretie A (2011) Africa needs climate data to fight disease. Nature 471: 440-442.

94. Paaijmans KP, Read AF, Thomas MB (2009) Understanding the link between malaria risk and climate. Proceedings of the National Academy of Sciences of the United States of America 106: 13844-13849.

95. Peterson I, Borrell LN, El-Sadr W, Teklehaimanot A (2009) Individual and household level factors associated with malaria incidence in a highland region of Ethiopia: a multilevel analysis. Am J Trop Med Hyg 80: 103-111.

96. Mauny F, Viel JF, Handschumacher P, Sellin B (2004) Multilevel modelling and malaria: a new method for an old disease. Int J Epidemiol 33: 1337-1344.