Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Alice E. Parish Date

A Spatial Analysis of Malaria and Related Health Outcomes

in Pregnant Women in Guatemala

By

Alice E. Parish

MSPH

Emory University

Rollins School of Public Health

Department of Biostatistics

_____ [Chair's Signature]

Lance A. Waller

_____ [Member's Signature]

Jodi Vanden Eng

A Spatial Analysis of Malaria and Related Health Outcomes in Pregnant Women in Guatemala

By

Alice E. Parish

B.S., Presbyterian College, 2012

B.A., Presbyterian College, 2012

MSPH, Emory University

Rollins School of Public Health

2014

Advisor: Lance A. Waller, Ph.D.

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 Science in Public Health

in Biostatistics

2014

A Spatial Analysis of Malaria and Related Health Outcomes

in Pregnant Women in Guatemala

By: Alice E. Parish

Abstract:

This paper focused on the spatial distribution of subclinical malaria cases and related health outcomes among pregnant women in Fray Bartolome de las Casas, Guatemala. Clinical data were gathered at a local health clinic during antenatal care visits; spatial data, including household points and routes, were gathered using GPS tracking devices. These data were linked, using a unique identification number, to test for spatial clustering of health outcomes and associations with physical access to care, with regards to distance. The binary health outcomes of interest were subclinical malaria (of any species), subclinical *P. vivax*, subclinical *P. falciparum*, anemia, severe anemia, delivery of a low birth weight infant, and premature delivery. It was found that there was no significant difference in distance from the health facility between the two possible outcomes for each of these factors when applying a t-test. Logistic regression was used to determine significant associations between distance, maternal age, number of previous gestations, and mother's weight for each of the binary health outcomes. Mother's weight was a significant predictor for subclinical *P. vivax*, low birth weight, and severe anemia (p-values = 0.019, 0.0001, 0.023, respectively) when controlling for all other factors. Maternal age is also significantly associated with the risk of subclinical *P. vivax* (p-value = 0.0355) when holding all other variables constant. Ripley's K function and Kulldorff spatial scan methods were applied to the same variables as before to assess global and local clustering. Ripley's K function, with edge correction, determined strong spatial clustering of cases at all distances within the study area. Three significant clusters were calculated with Kulldorff's scan method – two clusters of subclinical *P. falciparum* (p-values = 0.010, 0.011) and one cluster of anemia (p-value = 0.046). These results and further research will help reduce the risk of malaria in pregnancy and related health outcomes for women in this region of Guatemala through education, behavior change, and policy reform.

A Spatial Analysis of Malaria and Related Health Outcomes in Pregnant Women in Guatemala

By

Alice E. Parish

B.S., Presbyterian College, 2012

B.A., Presbyterian College, 2012

MSPH, Emory University

Rollins School of Public Health

2014

Advisor: Lance A. Waller, Ph.D.

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 Science in Public Health

in Biostatistics

2014

Acknowledgements

I want to thank the faculty, advisors, and staff of the Biostatistics and Bioinformatics Department at Rollins School of Public Health for their dedication to teaching and love of learning, and passing this on to me. I have gained so much knowledge and experience over the past two years and this thesis is an example of that. I would like to especially thank Dr. Lance Waller for his advice while writing my thesis.

Also a special thanks to the Centers for Disease Control and Prevention for the opportunity to work on this project and to Jodi Vanden Eng for her support and guidance.

A personal thank you goes to my family and friends for inspiring and motivating me to pursue a graduate degree, and for their encouragement along the way.

Table of Contents

Chapter 1: Introduction

Background on Malaria

Malaria is a parasitic disease that affects people worldwide. In 2012, there were an estimated 207 million clinical cases, 627,000 deaths, and over 3.3 billion people at risk of malaria ("Malaria: Malaria Facts, cdc.gov; "Malaria", who.int). There are four species of the malaria parasite that affect humans – *Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae,* and *Plasmodium ovale*. *P. vivax* accounts for approximately 74% of the malaria burden in the Americas, *P. falciparum* for 25%, and *P. malariae* for less than 0.4% ("Regional Strategic Plan for Malaria", 2006). *P. falciparum* is the most deadly form of malaria and is most common in Africa. In Central and South America *P. falciparum* and *P. vivax* are the two dominant malaria species.

Malaria is spread through the infected bite of the female mosquito, predominantly of the Anopheles genus. *Anopheles albimanus* and *Anopheles pseudopunctipennis* are two of the common vector species found Central America ("Malaria: Global Distribution…", cdc.gov). Malaria is most common in subtropical and tropical regions due to the favorability of these environments for the mosquito vector; it is uncommon in areas of high altitudes or colder climates where the mosquito vectors are absent or rare.

The parasite enters a person's bloodstream through the saliva of the mosquito. Once someone has been infected, it takes an average of 7 to 30 days to experience symptoms ("Malaria: Disease", cdc.gov). Symptoms can vary widely; some people remain asymptomatic, while others experience flu-like symptoms of fever, chills, sweats, headache, nausea, vomiting, and body aches. The most susceptible populations are children under five years of age, pregnant women, the immunocompromised, and travelers or migrants, who have little or no immunity to

malaria. More severe cases can lead to seizures, severe anemia, respiratory distress, low blood pressure, kidney failure, and death ("Malaria: Disease", cdc.gov).

The severity of health outcomes of malaria during pregnancy varies depending on several factors including level of acquired immunity, species of parasite, and number of previous births. In areas with high endemicity, where women have higher levels of acquired immunity, malaria infections during pregnancy are most likely to remain asymptomatic. However, the parasite can infect the placenta, and cause maternal anemia or low birth weight in the newborn, both of which increases the risk of infant mortality. These risks are most pronounced in first time mothers infected with *P. falciparum* ("Malaria in Pregnant Women", who.int).

In low transmission areas, such as Guatemala, women have a lower level of acquired immunity. This makes all pregnant women more susceptible to severe malaria, anemia, spontaneous abortion, stillbirth, prematurity, and low birth weight, regardless of number of previous pregnancies or parasite species (either *P. vivax* or *P. falciparum*) ("Malaria in Pregnant Women", who.int).

Background on the Study

This study is part of a larger project, PregVax, which is a multi-center descriptive effort researching the effects of malaria during pregnancy. Researchers have collected clinical data from Brazil, Colombia, Guatemala, India, and Papua New Guinea as well as spatial data from Brazil, Colombia, and Guatemala for the years 2009 to 2011.

Clinical data were collected at the time of enrollment (coinciding with the first antenatal clinic visit (ANC) made by the pregnant women), at each subsequent ANC visit, during any unscheduled sick visits, and at delivery if the delivery occurred at the health center. The data collected focused on symptoms associated with malaria such as anemia, fever, and presence of

malaria in a blood sample, regardless of symptoms. A portion of the collected blood samples were also tested for the presence of very low level, submicroscopic malaria, using polymerase chain reaction (PCR), also known as subclinical malaria.

This paper will focus on the clinical and spatial data collected in Fray Bartolome de las Casas, Guatemala. Fray Bartolome has close to 55,000 inhabitants and experiences year-round, low-level transmission of malaria.

Problem Statement

Little is known about factors influencing the risk of malaria during pregnancy, such as whether access to healthcare impacts risk. There is also a lack of knowledge about potential differences in clinical outcomes and severity between *P. falciparum* and *P. vivax* in association with pregnancy, particularly in areas of low prevalence including the Americas. However, such information is needed to make informed decisions for policy reform, resource distribution, and interventions targeted towards pregnant women in malaria risk areas.

Purpose Statement

Due to the importance of understanding malaria risk in pregnant women, we aim to answer the following questions about malaria in pregnancy for the Guatemala study site:

- Is there clustering of malaria, or other clinical health outcomes including submicroscopic malaria, anemia, low birth weight, or prematurity, in the study area?
- Are there different clustering patterns between *P. falciparum* and *P. vivax*, the two most common species of malaria in Guatemala?
- Is the distance from a health care facility associated with malaria and other health outcomes, such as malaria, subclinical malaria, anemia, prematurity, and low birth weight? Is there an association between distance from a health care facility and

location of delivery (health care center versus home birth)?

Significance Statement

The results from this paper provide increased understanding of risk patterns in order to better protect at-risk pregnant women from malaria through informed policy change, better distribution of resources, and education.

Chapter 2: Review of the Literature

We briefly review the associated literature with several specific goals in mind. First, we summarize past research regarding distance and access to health care, factors influencing access, and outcomes impacted by access to health care. Second, we review how Geographic Information Systems (GIS) have been used to study and measure malaria risk. We limit our search to literature published in English within the last ten years.

Measuring Physical Access to Healthcare

Access to care can mean several things including physical access and financial access. Physical access is the physical ability for a person to get to a healthcare facility. It can be measured in distance or time. Financial access assesses the affordability of healthcare for an individual. The papers reviewed all focus on physical access measured in distance from a healthcare facility.

Noor et al. (2003) examined the difference in access to care in rural populations versus urban populations in Kenya. Using government health facilities only, data about potential use, and straight line distances, Noor et al. (2003) estimated actual use for each facility. The researchers found a general reduction in the usage of government health facilities as distance from the facilities increased. In general, urban populations have good physical access to care, living within 5 kilometers of a facility on average, unlike rural populations.

In another study, which took place among the Yanomami people in Amazonian Venezuela and Brazil, Grenfell et al. (2008) examined whether varying levels of access to healthcare affected health outcomes, including malaria and anemia. The Yanomami communities were divided into those that had consistent access to health care through a permanent health post, and those with intermittent access to care from a mobile medical team. Grenfell et al. (2008) found a significant difference between the two types of communities in

terms of malaria, anemia, splenomegaly, fever, and diarrhea. There was a strong association (p-value < 0.001) between malaria prevalence and access. A strong association was also found between low hemoglobin and access to care when controlling for age, gender, malaria, splenomegaly, and clustering at the community level.

Baker (2008) aimed to create a rapid assessment of physical access to care in a mountainous area of Honduras. Using a GPS device to map routes from an NGO (Non-Governmental Organization) healthcare facility to a random selection of patients' houses, they took the "cookie crumb trail" of GPS points and applied a travel cost to the routes based on physical characteristics of the route, such as road versus footpath, so that a travel time could be estimated. They aimed to show that this method is a good way to determine the catchment area for a health facility. The study area was a rural, low-income community comparable to many other small communities in Latin America. They determined that there is an almost 50% loss of attendance to the health facility once distance surpasses 4 kilometers.

In Yemen, Al-Taiar (2010) measured access to healthcare by calculating the driving distance and time from a participant's house to the facility, in both urban and rural areas. The author used GPS (global positioning system) coordinates to take note of the start and finish locations (residence and facility locations). These distances provided a straight-line distance to determine if physical access to care was associated with vaccination status in children. Al-Taiar (2010) found that there was a strong correlation among driving distance, straight-line distance, and driving time (p-value < 0.001) for each comparison. The results also reveal a strong correlation between each distance measurement and vaccination status in children.

Using GIS to Analyze Malaria Risk

Dogan et al. (2010) conducted a spatiotemporal study in Turkey examining

environmental determinants of malaria over the course of 34 years, 1975 to 2008. They found that over the decades, high malaria areas shifted from the east Mediterranean region to the southeast Anatolia region, which reflects the shift in human activities in Turkey over the years. Dogan et al. (2010) also found that latitude, minimum temperature, distance to seas, and elevation had significant impacts on malaria prevalence.

De Oliveira (2013) analyzed data of malaria incidence in a rural Amazonian village in Brazil originally collected in 2005. The authors used logistic regression to determine which environmental factors were associated with malaria risk and to predict relative likelihood of disease infection. GIS provided distance measurements between households and various local environmental factors such as gold mining sites and the closest body of water. They found that the highest malaria risk occurred in the southern region of their study area, which contained gold mining sites, high land usage, high levels of soil and vegetation humidity, and low vegetation density.

GIS can also be used to determine the location of hotspots of malaria risk. Nath et al. (2013) conducted a spatiotemporal analysis in Northern India from 2006 to 2009 and mapped the annual parasitic index (API), slide positivity rate, annual blood examination rate, and *P. falciparum* percentage for each health sub-center in the Sonitpur district of Assam, India. The maps were overlaid and analyzed to define high risk areas for malaria. This information can be very useful in prioritizing areas for malaria control efforts. Coulibaly (2013) conducted a similar study in Bandiagara, Mali from 2009 to 2010. Spatial and temporal data included household locations, clinical outcomes, and rainfall. The researchers used GIS and SaTScan™ (Kulldorff, 2005), a program that detects significant geographic clustering, to map clusters of high and low malaria risk.

Subclinical Malaria

Lindblade et al (2013) reviewed what is currently known about subclinical malaria infection. There is no standard definition of asymptomatic parasitemia, but there must be detection of the parasite, either in a sexual or asexual stage, as well as an absence of clinical symptoms of malaria during a specified time frame. The time frame can range from the moment the patient is seen up to 60 day after the visit. If there is no follow up, a patient with parasite detection, but lack of symptoms, could be misclassified as subclinical malaria case, when in fact the patient may be in the incubation phase of malaria, which occurs right after infection and right before symptoms show. However, some argue that this scenario would not be very common since the incubation stage for *P. falciparum* is 1 day, and for *P. vivax* it is 4 days. Some definitions outline clinical symptoms of malaria as all symptoms associated with the disease, while others just focus on the most common symptom of fever. In general, asymptomatic parasitemia only includes the presence of malaria in the blood stream, not in the liver, due to the fact that *P. vivax* can remain dormant in the liver for months or years after infection. Also some people define clinical malaria based on parasite density thresholds, rather than presence or absence of symptoms. In this case, patients with a fever but low parasitemia would be classified a subclinical malaria case.

Malaria in Pregnancy

A case report of congenital malaria was published based on findings from the PregVax study. Castellanos et al. (2012) reported that a 21 day old baby was admitted to the hospital in Fray Bartolome de las Casas, Guatemala with a fever, hepatosplenomegaly, anemia, and was positive for *P. vivax*. The infant was treated with the recommended antimalarial medication. The mother had no previous history of malaria during pregnancy, but reported a fever at the time

of delivery and tested positive for *P. vivax*. Further analysis showed that the same parasite genotype was detected in the mother's peripheral blood, placental blood, cord blood, and the infant's peripheral blood, which reinforces the theory that this was a case of congenital malaria, that is, a mother-to-child transmission of malaria.

Souza et al. (2013) studied the effect of malaria on the placenta, specifically in the case of *P. vivax.* Most of what is known about malaria in pregnancy is based on findings associated with *P. falciparum*. In this study, placentas from malaria-exposed women were examined for signs of the parasite and ten other histological parameters. The study took place in Brazil and women were diagnosed with malaria using microscopy, which may overlook asymptomatic cases. In placentas from women exposed to *P. vivax*, they found lesions, syncytial knotting, placental barrier thickness, and mononuclear cells were increased compared to placentas from non-exposed women.

A study in Colombia also analyzed clinical outcomes of malaria in pregnancy. Piñeros et al. (2013) aimed to determine if there were clinical differences between *P. falciparum* and *P. vivax* infections in a low-transmission endemic area. All of the women in the study had malaria infections; the severity of the infection ranged from asymptomatic (10.2%) to severe (9.0%). The most common complications were hepatic dysfunction and spontaneous bleeding, found in greater than 70% of severe malaria cases; no significant difference was found between malaria species. Organ-specific complications occurred frequently in *P. vivax* infections, but not significantly more than in *P. falciparum* infections.

Chapter 3: Methodology

Data Collection

Recruitment of 2,000 pregnant women began in 2009 at a chosen health facility that provided antenatal care. The facility was chosen based on the high malaria prevalence in the area and for the high rate of ANC attendance. The study was approved by the Institutional Review Boards (IRBs) at the Centers for Disease Control and Prevention (CDC) and Universidad del Valle de Guatemala (UVG). Pregnant women who qualified were recruited to enroll in the PregVax clinical study and an ancillary spatial study. Women were enrolled in the clinical study first; only women consenting to the clinical study were eligible for enrollment in the spatial study. Each woman who provided informed consent received a unique study identification number; the same number was used in the clinical study as in the spatial study so that data could be linked at a later time.

At the time of enrollment for the spatial study, women provided directions from the clinic to their household on the recruitment form. Using these instructions, researchers travelled to the household during an agreed up window of dates. A handheld PDA device with GPS capabilities was used to track the route. A GPS coordinate would be saved every 50 meters or 15 seconds, whichever came first. This trail of points was subsequently connected to form a complete route from the healthcare facility to the study participant's household. To simplify logistics and reduce cost, multiple participants living in the same area were visited in a single trip. A primary route from the health facility was recorded then a secondary route would connect the primary route to a household. Route data was taken twice, once on the way to the household, and again on the way back to the clinic to ensure accuracy. At the domicile, three GPS points were taken, and the average of the three was used as the location of the household. A short questionnaire

was also given on household visits, which asked about basic malaria-associated behaviors such as insecticide treated bed net (ITN) usage. The house point and route data were mapped in ArcGIS 10.2 (ESRI, 2013) as can be seen in Map 1.

Clinical data were collected at each visit to the health clinic (Table 1). Blood samples were taken to determine presence of malaria and level of anemia. The study implemented active and passive detection of infection. Birth outcomes, such as complications, abortion, prematurity, and low birth weight, were recorded at delivery. For this study, asymptomatic malaria (subclinical malaria) was defined as presence of asexual parasite forms of any *Plasmodium* species on thick or thin blood smears, or a positive PCR test, in the absence of clinical symptoms. Anemia was defined as hemoglobin levels less than 11 g/dL and severe anemia is hemoglobin levels less than 7 g/dL. Low birth weight was defined as a birth weight less than 2,500 grams. Premature delivery was any birth with a gestational age less than 37 weeks.

Statistical Analysis– Access to Healthcare

A t-test is a statistical method used to determine if there are significant differences in the mean value of an attribute between two groups. This can be used to compare mean distances from a health care facility to a residence between individuals with different binary health outcomes such as presence/absence of subclinical malaria (any malaria, *P. vivax,* and *P. falciparum*), place of delivery (at home or at a health center), presence of anemia, and premature birth. Here we define distances between the house point and the health center as the length of the route determined by GPS tracking, rather than as a Euclidean ("as the crow flies") distance.

A logistic regression model measures the association between continuous or categorical predictor variables and a binary outcome. For this analysis, the health outcomes listed above will be modelled using distance from the health facility, mother's weight, age, and number of

previous gestations as predictors.

Statistical Analysis– Clustering

A cluster can be described as a pattern in which locations or events tend to occur in close proximity to each other more frequently than expected (Waller and Gotway, p118). There are several tests that can be used to determine if there is significant global, local, or focal clustering of spatial point data. Each of the following tests uses a homogeneous Poisson point process with a null hypothesis of complete spatial randomness, that is, no uniformity or clustering.

Ripley's K function can detect global clustering (or uniformity) of individual level case data (Ripley, 1977). The K function can identify patterns such as distances where heterogeneity begins, where clustering becomes significant, and where maximum clustering is observed. It measures how many additional events (or cases) occur within a given distance of a randomly chosen event. The K function is defined as the expected number of additional events within a distance, *h*, of a randomly chosen event divided by a constant intensity, λ. It is estimated as

$$\widehat{K}_{ec}(h) = \hat{\lambda}^{-1} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} w_{ij}^{-1} \mathrm{I}(d(i,j)) \qquad (1)$$

where *N* is the number of events, $\hat{\lambda} = (\frac{\text{number of events in } A}{|A|})$, *A* is the study area, |*A*| is the area of *A*, *d(i, j)* is the Euclidean distance between points *i* and *j*, and $\mathrm{I}(d(i,j)) = 1$ if *d(i, j)* < *h*, for a given distance *h*, and is 0 otherwise. This estimate uses an edge correction term, w_{ij}, which weights points with "a proportion of the circumference of the circle centered at event *i* with radius *d(i, j)*" (Waller and Gotway, 138). The weight has a value of 1 if the distance between case *i* and *j* is less than the distance between *i* and the edge of the study area. (Waller and Gotway, 137-8)

It is useful to plot distance versus *L(h)* to determine at what distances events are clustered or random. The L function is estimated as

$$\hat{L}(h) = \sqrt{\frac{\hat{K}_{ec}(h)}{\pi}} \quad (2)$$

The null hypothesis of interest is that the distribution of cases follows a spatial Poisson point process, where $L(h) = h$. The alternative of interest is that the cases are clustered at some distances, that is, $L(h) > h$ for some h (Waller and Gotway, 139). We will use this test to calculate if there is global clustering at some distances for the binary health outcomes subclinical malaria, anemia, low birth weight, and prematurity.

Another way to test for local clustering is to use Kulldorff spatial scan statistic (1997). This technique uses varying window sizes at each point to determine if there is a significantly higher incidence of cases in an area while accounting for possible clustering or heterogeneity in the background population density. The scan statistic also identifies the location of the most likely clusters. The statistic is based on the log likelihood ratio, constructed of observed and expected number of cases, and tests significance via Monte Carlo simulation. Using SaTScan™ (Kulldorff, 2005) we will test if there is local clustering of the previously mentioned health outcomes. Any significant clusters will be mapped to visualize spatial patterns.

Chapter 4: Results

Descriptive Statistics

A total of 2,009 women were enrolled in the clinical study of which sixteen withdrew. Spatial data were collected for 1,554 households. Eighty percent of women recruited in the clinical study were of Mayan ethnicity, 15% identified as Latino, and 5% did not specify. The average age of the women was 24 years old, with the youngest reported to be 12 and the oldest 46 years old. The average weight at enrollment was 129 pounds, ranging from 73 to 234 pounds. Women had an average of three previous pregnancies, ranging from 0 to 14. Table 2 summarizes the previous birth history collected from the participants, including number of previous live births, neonatal deaths, premature births, spontaneous abortions, and still births.

There were only 5 cases of clinical malaria, all of which were of the species *P. vivax*. Since this outcome was rare, no statistical tests were performed. A random sample consisting of 1,242 blood samples (20% of the total) was tested for subclinical malaria. These blood samples came from several sources (maternal peripheral blood, placental blood, and cord blood) and several time points (recruitment, sick visits, and delivery); it was possible for two samples from the same woman, from a different source or time point, to be tested. 3.6% of the blood samples were positive for *P. falciparum* and 11.4% were positive for *P. vivax* (Table 3). Seventeen samples (1.4%) were positive for both *P. vivax* and *P. falciparum* (Table 3). There were no positive samples for *P. malariae* or *P. ovale*. Of the 1,242 samples, 544 (43.8%) were linked to household points using a unique subject identification number.

In addition to malaria outcomes, related pregnancy and infant outcomes were evaluated. Of the 2,009 women recruited, only 1,032 (51.3%) were followed up at delivery, with 1,011 newborn records. Almost 7% of the infants were low birth weight and there were forty-two

(4.2%) recorded premature births. Of the 1,032 deliveries recorded, 990 (95.9%) took place at a hospital or health care facility. Anemia data was collected throughout pregnancy; 337 (32%) of 1,025 tested women experienced anemia during pregnancy, seven of these women experienced severe anemia. (Table 3)

Statistical Analyses – t-Test, Logistic Regression

Eight binary health outcomes were tested using the t-test to determine if there was a significant difference in average distance from the health facility between the two outcomes (Table 4). No significant differences were found for any of the variables using a significance level 0.05. However, the data did not completely follow a normal distribution, with a higher frequency of short distances. Figure 1 shows the Q-Q plot for any subclinical malaria; the Q-Q plots for the other seven variables were similar.

A logistic regression was performed for each binary health outcome, testing if distance, age, number of previous gestations, and mother's weight were significant predictors (Table 5). These factors were not significant at the 0.05 level for any of subclinical malaria, subclinical *P. falciparum*, co-infection, prematurity, or anemia.

Both maternal age and weight were significant predictors for subclinical *P. vivax* when controlling for the other factors (p-value = 0.04, 0.02 respectively). Thus, for each one-year increase in age, holding all other variables constant, there is a 0.945 (95% CI: 0.896, 0.996) multiplicative change in odds of having subclinical *P. vivax*. For every one pound increase in weight, there is a 1.013 (95% CI: 1.002, 1.023) multiplicative increase in the odds of having subclinical *P. vivax*, controlling for all other factors.

The mother's weight was a significant protective predictor for low birth weight and severe anemia, when holding all other variables constant. For every one pound decrease in

weight there was a 1.031 (95% CI: 1.015, 1.047) increase in the odds of having a low birth weight infant. The odds of having severe anemia is increased by 1.064 (95% CI: 1.009, 1.123) for every one pound decrease in weight.

Spatial Analyses

Ripley's K function was applied to locations for each of the binary health outcomes. It is most trustworthy when the number of cases is greater than thirty, so severe anemia and place of birth were not tested. This approach requires a boundary of the study area; a minimum bounding rectangle was calculated based on all mapped house points (Map 13). This boundary was used for all health outcomes. An edge correction was applied to each calculation to account for points close to the edge of the rectangle. Figures 2 through 7 indicate that all health outcomes were strongly clustered for all distances in the study area. A confidence envelope was determined based on 99 permutations. Each confidence envelope captured the expected distribution of cases based on complete spatial randomness in the rectangular study area.

Kulldorff spatial scan statistic, following a Bernoulli probability model, was implemented to determine if there were clusters of high risk- that is, if there are areas with a significantly higher concentration of cases than expected under uniform assignment. This method was applied over the entire study area, as well as focused in on the urban center near the healthcare facility, and in just the rural areas. There was a significant high risk cluster of subclinical *P. falciparum*, which encompassed 284 participants. The area had 16 subclinical cases, leading to a relative risk 5.70 times more cases than expected (Table 6, Map 10). When the data were divided into rural and urban, a significant cluster for subclinical *P. falciparum* was detected in the southeast rural region (Table 6, Map 11). One significant cluster was calculated for high risk of anemia over the total study area also in the southeast region of the study area (Table 6, Map 12). There were only

seven cases of severe anemia that were connected to women from spatially identified households.

Chapter 5: Discussion

Conclusions

The occurrence of subclinical malaria, of any species, was fairly spread out with no apparent visible patterns either for the entire study area or in the urban center near the healthcare center (Map 2). When the malaria is divided by species there does seem to be some difference in the areas where each is occurring. *P. falciparum* seems to occur mainly in the southeast part of Fray Bartolome (Map 3), whereas *P. vivax* is spread throughout the study area (Map 4). This pattern is supported by the Kulldorff spatial scan statistic. No significant clusters were found for any subclinical malaria or specifically for *P. vivax*, but there was one significant cluster for high risk of *P. falciparum*, which highlights the southeast region (Map 10). The difference in incidence for these two species in the area is important; there could be environmental factors influencing the difference, or behavioral factors that are or are not taking place to create the divide.

Low birth weight and prematurity appear to be randomly distributed throughout the study area and within the urban center (Maps 6 and 7). The lack of significant clustering is supported by the Kulldorff spatial scan statistics (Table 6). At first glance, it seems like anemia is randomly spread through the entire study area including within the urban center (Map 8), however the Kulldorff spatial scan detected a cluster of high risk in the southeast region. No significant difference was found in the average distance from the health facility for anemia or severe anemia when applying the t-test (Table 4). While there were only seven cases of severe anemia, it is interesting to note that none of them occurred in the urban center (Map 9).

The significant global clustering for all of the variables tested with Ripley's K function is not surprising as the calculated study area boundary is not the most accurate representation of the

study area. The house points are clustered near the health center and there is a lot of empty space in the outer regions of the boundary especially on the western edge (Map 13). While an edge correction was applied, this empty space might contribute to a false positive global clustering at all distances. If a better study area boundary was determined that followed the edges of the study area and minimized empty space, a more accurate Ripley's K test could be performed. The function used in this analysis based distances on Euclidean distance between two points. Better precision for complete spatial randomness might be achieved if network distances were considered instead.

Strengths

One of the biggest strengths of this study is how distance was measured. Many studies use Euclidean (straight line) distance to measure between two points. However, this is not always an accurate or precise way to measure physical access to care. How far a person has to travel to a health facility is dependent on roads and road quality. A straight line distance might indicate that a person is close to a hospital, but in reality they might have a much longer distance to travel, depending upon the presence of obstacles such as a forest or mountain. This study used GPS points to create a more accurate representation of the routes taken by the women to access the health facility. Greater accuracy in measuring distance leads to more accurate and trustworthy results.

Another strength of this study is the use of PCR to detect subclinical malaria. The region in the study has low endemicity, so it was unlikely that there would be many clinical cases of malaria. However, by testing for subclinical malaria, it is possible to determine what areas still have a high prevalence of asymptomatic cases. Identifying these areas is imperative in the pursuit of elimination of malaria.

Limitations

The PCR data used in this analysis is preliminary data. There were a total of 6,324 blood samples taken. 20% (1,340) were tested for subclinical infections using PCR. Of these, 393 were missing study identification numbers, and were unable to be linked to a household for spatial analysis. There were multiple PCR samples for some women since samples can come from the mother (during recruitment, a sick visit, and at delivery), the cord blood, or the placenta. The assumption was made that if a woman had malaria at recruitment and then again at delivery (two separate time points), that these were two separate infections, and that she was treated and cleared of malaria in between the two times, per protocol.

One limitation of the study involves time under observation. All of the analyses done were purely spatial. However, women contributed varying amounts of time to the study depending on gestation age when recruited. In other words, a woman entering the study during first trimester would contribute more observation time than a woman entering during her third trimester. Also, the study encompassed a two year period. While endemic, malaria is known to be seasonal, so calendar time could impact the risk of malaria for pregnant women and merits future follow up.

Since the clinical data were collected at varying times, there is varying loss of follow up for each time point, and thus for each health outcome analyzed. Also, clinical data and spatial data were collected separately of one another, so there is some loss of information between the two datasets if study identification numbers cannot be matched. Table 7 shows the loss of information for each health outcome after clinical and spatial data were joined.

Birth outcomes were difficult to assess since only eight houses in our mapped routes were associated with home deliveries. Map 5 shows the locations of reported place of delivery. Only

one home birth took place within the urban center, and all of the other reported births occurring in a hospital or health facility are concentrated around the health facility. Looking at the map, it appears that the rural areas on the outskirts of the study area have little or no data collected for this variable which likely indicates that there was a home birth (no follow up was conducted for home deliveries). The lack of data creates a bias in the birth outcome variables low birth weight and prematurity. It also biases the random selection of blood samples for PCR testing, as no samples at delivery were taken for the women who had home births, skewing the sample pool.

Implications and Recommendations

The most striking finding in this study is the difference between *P. falciparum* and *P. vivax*. *P. vivax* is uniformly spread throughout the study region, as demonstrated by the cluster analyses and maps. There were more clinical and subclinical infections of this species in the population. *P. falciparum*, on the other hand, is not as prevalent as *P. vivax*, but is more concentrated in a specific area. Future research can explore why this species is focused in the area. Environmental and behavioral factors should be taken into consideration. Many questions are of interest: are there more water sources in the area, a different species of mosquito, or is the elevation or average rain fall significantly different from the rest of the study area? In this area, are the households more open, is there a lack of insecticide treated bed nets (ITNs)? Time could also play an important role in this difference between the two species. Was the cluster of *P. falciparum* an isolated epidemic that occurred during a finite window of time in the two year study period, or is it more endemic and spread out? Another interesting result involves anemia. High risk of anemia was clustered in the same area as asymptomatic *P. falciparum*. As anemia is associated with malaria, it is likely that the cluster of anemia in this area is due to the cluster of subclinical *P. falciparum.* However, further investigation should examine if there are other local

factors, such as low iron intake, that are the cause of the low hemoglobin levels.

Another important result to follow up with is the lack of home birth data. A follow up of the health outcomes and possible asymptomatic malaria from these births is necessary to better represent the population of women who give birth at home.

These results increase our knowledge of malaria risk patterns in pregnant women in Fray Bartolome de las Casas, Guatemala. While asymptomatic malaria is spread throughout the region, there are differences between malaria species. The region in the southeast has a higher risk of *P. falciparum* subclinical infections than other regions. Further research into possible factors that could be associated with the cluster of malaria is recommended so that education, policy reform, or an intervention can be implemented to reduce the risk of malaria for the pregnant women.

.

Works Cited

Al-Taiar, A., et al. "Physical Accessibility and Utilization of Health Services in Yemen." *International Journal of Health Geographics* 9.1 (2010): 38. Print.

Baker, J. B., et al. "Rapid Assessment of Access to Primary Care in Remote Parts of the Developing World." *Field Methods* 20.3 (2008): 296-309. Print.

Castellanos, M. E., et al. "Plasmodium Vivax Congenital Malaria in an Area of Very Low Endemicity in Guatemala: Implications for Clinical and Epidemiological Surveillance in a Malaria Elimination Context." *Malaria Journal* 11.1 (2012): 411. Print.

Coulibaly, D., et al. "Spatio-temporal Analysis of Malaria within a Transmission Season in Bandiagara, Mali." *Malaria Journal* 12.82 (2013) Print.

De Oliveira, E. C., et al. "Geographic Information Systems and Logistic Regression for High-resolution Malaria Risk Mapping in a Rural Settlement of the Southern Brazilian Amazon." *Malaria Journal* 12.420 (2013) Print.

Dogan, H. M., et al. "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* 104.11 (2010): 726-32. Print.

ESRI (Environmental Systems Resource Institute). 2013. ArcMap 10.2. Redlands, CA. Computer software.

Grenfell, P., et al. "Anaemia and Malaria in Yanomami Communities with Differing Access to Healthcare." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102.7 (2008): 645-52. Print.

Kulldorff, M. "A Spatial Scan Statistic." *Communications in Statistics - Theory and Methods* 26.6 (1997): 1481-496. Print.

Lindblade, K. A., et al. "The Silent Threat: Asymptomatic Parasitemia and Malaria Transmission." *Expert Review of Anti-infective Therapy* 11.6 (2013): 623-39. Print.

"Malaria in Pregnant Women." *WHO*. 6 Mar. 2013. Web. 04 Mar. 2014. <http://www.who.int/malaria/areas/high_risk_groups/pregnancy/en/>.

"Malaria." *WHO*. World Health Organization, Dec. 2013. Web. 15 Jan. 2014. <http://www.who.int/mediacentre/factsheets/fs094/en/ >.

"Malaria: Disease." *Centers for Disease Control and Prevention*. Centers for Disease Control and Prevention, 08 Feb. 2010. Web. 12 Jan. 2014. <http://www.cdc.gov/malaria/about/disease.html>.

"Malaria: Global Distribution (Robinson Projection) of Dominant or Potentially Important Malaria Vectors." *Centers for Disease Control and Prevention*. Centers for Disease Control and Prevention, 08 Feb. 2010. Web. 15 Jan. 2014. <http://www.cdc.gov/malaria/about/biology/mosquitoes/map.html>.

"Malaria: Malaria Facts." *Centers for Disease Control and Prevention*. Centers for Disease Control and Prevention, 19 Sept. 2012. Web. 15 Jan. 2014. <http://www.cdc.gov/malaria/about/facts.html>.

Nath, M. J., et al. "Prioritizing Areas for Malaria Control Using Geographical Information System in Sonitpur District, Assam, India." *Public Health* 127.6 (2013): 572-78. Print.

Noor, A. M., et al. "Defining Equity in Physical Access to Clinical Services Using Geographical Information Systems as Part of Malaria Planning and Monitoring in Kenya." *Tropical Medicine and International Health* 8.10 (2003): 917-26. Print.

Piñeros, J. G., et al. "Maternal Clinical Findings in Malaria in Pregnancy in a Region of Northwestern Colombia." *American Journal of Tropical Medicine and Hygiene* 89.3

(2013): 520-26. Print.

Regional Strategic Plan for Malaria in the Americas 2006-2010. Publication. Pan American Health Organization, 2006. Web.

Ripley, B. D. "Modeling Spatial Patterns." *Journal of the Royal Statistical Society* B 39.2 (1977): 172-212. Print.

Souza, R. M., et al. "Placental Histopathological Changes Associated with Plasmodium Vivax Infection during Pregnancy." *PLoS Neglected Tropical Diseases* 7.2 (2013): E2071. Print.

Waller, L. A., and Carol A. Gotway. *Applied Spatial Statistics for Public Health Data*. Hoboken, NJ: John Wiley & Sons, 2004. Print.

Appendix A: Tables

Table 1: Time points and contacts with study participants

Time points	1st ANC visit	2nd ANC Visit	3rd ANC visit	Delivery
All pregnant women attending ANC	1. Informed consent and enrolment 2. Demographic data 3. Obstetric and clinical data 4. ADI 5. Hb 6. Venous blood sampling in a *sub-sample* of women 7. Venous blood sampling in *P. vivax* infected women*	1. ADI 2. Obstetric and clinical data 3. Blood sampling in *P. vivax* infected women*	1. ADI 2. Obstetric and clinical data 3. Blood sampling in *P. vivax* infected women*	1. Foetal and maternal outcomes 2. ADI 3. Hb 4. Maternal venous blood and cord blood sampling in *P. vivax* infected women* and in the *sub-sample* of women 5. Placental samples in all women** 6. Newborn (smears and filter papers)
All sick women attending hospital	Passive Case Detection (clinical information, blood smears and/or RDT, filter papers, Hb) Blood sampling if *P. vivax* infected women *			

- ANC: antenatal clinic
- Blood sampling: 5mL peripheral blood sample for immunological and adhesion studies. At delivery, 5mL of cord blood will also be obtained for immunology. At some study sites, 2-3mL of placental blood may be collected.
- ***P. vivax* infected women will be defined as the women with *P. vivax* infection, with or without other *Plasmodium* co-infections**
- ADI: active detection of infection. Blood smears for detection of parasitemia and filter papers for PCR (if parasitemia positive, provide antimalarial treatment). In places where quick parasitemia assessment is not be possible, the ADI will be done by RDT, and a blood smear will also be collected in these cases
- Obstetric and clinical data include gestational age, parity, axillary temperature and signs/symptoms suggestive of a malaria clinical episode (see definitions section)
- **** Placental samples (impression smears and filter papers) will be collected straight from ALL women at delivery. Placental biopsies will be collected ONLY from women who tested positive in the peripheral blood for either *P. vivax, P. falciparum* or both, whether at recruitment, ANC visits, PCD visits or delivery, and in the *sub-sample* of women (identified at the time of enrolment).**
- Maternal outcomes include anaemia, severe complications and abortion
- Foetal outcomes include birth weight, prematurity, stillbirth and congenital infection
- Antenatal care local calendars will be followed

Table 2: Descriptive Statistics (Continuous)

Variable*	N	Missing	Mean	Min	Max
Maternal age (years)	2001	8	24.38	12	46
Maternal weight (lbs.)	2006	3	129.11	73.00	234
Number previous gestations	2004	5	2.99	0	14
Number previous live births	1208	801	2.95	1	12
Number previous neonatal deaths	116	1893	1.15	1	3
Number previous premature deliveries	4	2005	1.00	1	1
Number previous spontaneous abortions	128	1881	1.15	1	4
Number previous still births	114	1895	1.25	1	4
Hemoglobin	930	95	11.4	4.6	20.8

*All variables were collected at enrollment

Table 3: Descriptive Statistics (Categorical)

Variable	N	Missing	Categories	Frequency	Percent
Any clinical malaria	6324	0	Yes	5	0.1
			No	6319	99.9
Clinical *P. vivax*	6324	0	Positive	5	0.1
			Negative	6319	99.9
Any subclinical malaria	1242	98	Positive	169	13.6
			Negative	1073	86.4
Subclinical *P. vivax*	1242	98	Positive	141	11.4
			Negative	1101	88.6
Subclinical *P. falciparum*	1242	98	Positive	45	3.6
			Negative	1197	96.4
Subclinical *P. vivax* and *P. falciparum*	1242	98	Yes	17	1.4
			No	1225	98.6
Anemia (levels)	1025	0	No anemia	680	66.3
			Moderate	330	32.2
			Severe	7	0.7
			Nd	8	0.8
Anemia (n/y)	1025	0	No	680	66.3
			Yes	337	32.9
			Nd	9	0.9
Race	2009	0	Maya (Quechua)	1624	80.8
			Latino	304	15.1
			Nd	81	4.0
Low birth weight	1032	0	Yes	70	6.8
			No	928	89.9
			Nd	34	3.3
Premature	1011	0	Yes	42	4.2
			No	906	89.6
			Nd	63	6.2

Variable	N	Missing	Categories	Frequency	Percent
Sex of baby	1011	0	Female	475	47.0
			Male	529	52.3
			Nd	7	0.7
Place of delivery	1032	0	Hospital/health facility	990	95.9
			Home	10	1.0
			Other	11	1.1
			Nd	21	2.0

Nd – not determined

Table 4: t-tests

Variable	Difference in Mean Distance (meters)	t Statistic	P-value
Any subclinical malaria (No-Yes)	1692.8	1.16	0.25
Subclinical *P. vivax* (No-Yes)	2024.9	1.31	0.19
Subclinical *P. falciparum* (No-Yes)	893.6	0.33	0.74
Co-infection (No-Yes)	2630.1	0.66	0.51
Low Birth Weight (No-Yes)	-599.8	-0.33	0.74
Premature (No-Yes)	1863.0	0.77	0.44
Anemia (No-Yes)	1166.9	1.12	0.26
Severe Anemia (No-Yes)	-4517.9	-0.86	0.39

Table 5: Logistic Regression

$$logit(y) = \beta_0 + \beta_1(distance) + \beta_2(age) + \beta_3(no.\ previous\ gestations) + \beta_4(mother's\ weight)$$

Parameter	Estimate	Standard Error	OR Point Estimate	95% CI	Wald Chi-Square	P-value
y = any subclinical malaria						
Distance (total length)	-8.24E-6	7.597E-6	1.00	(1.00, 1.00)	1.1778	0.28
Age	-0.0487	0.0250	0.95	(0.91, 1.00)	3.7865	0.05
Number of previous gestations	0.1098	0.0638	1.12	(0.99, 1.27)	2.9641	0.09
Weight	0.00951	0.00507	1.01	(1.00, 1.02)	3.5107	0.06
y = subclinical *P. vivax*						
Distance (total length)	-9.36E-6	8.203E-6	1.00	(1.00, 1.00)	1.3014	0.25
Age	-0.0569	0.0271	**0.95**	**(0.90, 1.00)**	4.4202	**0.04**
Number of previous gestations	0.0894	0.0702	1.09	(0.95, 1.26)	1.6250	0.20
Weight	0.0125	0.00536	**1.01**	**(1.00, 1.02)**	5.4868	**0.02**

Parameter	Estimate	Standard Error	OR Point Estimate	95% CI	Wald Chi-Square	P-value
y = subclinical *P. falciparum*						
Distance (total length)	-5.33E-6	0.000014	1.00	(1.00, 1.00)	0.1485	0.70
Age	-0.0262	0.0451	0.97	(0.89, 1.06)	0.3383	0.56
Number of previous gestations	0.1041	0.1125	1.11	(0.89, 1.38)	0.8565	0.35
Weight	-0.00043	0.00970	1.00	(0.98, 1.02)	0.0020	0.96
y = co-infection						
Distance (total length)	-0.00001	0.000022	1.00	(1.00, 1.00)	0.2946	0.59
Age	-0.0629	0.0709	0.94	(0.82, 1.08)	0.7868	0.38
Number of previous gestations	-0.0508	0.2171	0.95	(0.62, 1.46)	0.0548	0.82
Weight	0.0114	0.0137	1.01	(0.99, 1.04)	0.6937	0.40
y = low birth weight						
Distance (total length)	6.033E-6	9.577E-6	1.00	(1.00, 1.00)	0.3968	0.53
Age	0.0400	0.0303	1.04	(0.98, 1.10)	1.7429	0.19
Number of previous gestations	-0.1486	0.0913	0.86	(0.72, 1.03)	2.6499	0.10
Weight	-0.0304	0.00799	**0.97**	**(0.96, 0.99)**	14.4474	**0.0001**
y = premature						
Distance (total length)	-0.00002	0.000014	1.00	(1.00, 1.00)	1.4432	0.23
Age	-0.00976	0.0428	0.99	(0.91, 1.08)	0.0520	0.82
Number of previous gestations	-0.0598	0.1241	0.94	(0.74, 1.20)	0.2320	0.63
Weight	0.00211	0.00829	1.00	(0.99, 1.02)	0.0651	0.80
y = anemia						
Distance (total length)	-7.17E-6	5.622E-6	1.00	(1.00, 1.00)	1.6282	0.20
Age	-0.00236	0.0183	1.00	(0.96, 1.03)	0.0166	0.90
Number of previous gestations	0.0631	0.0484	1.07	(0.97, 1.17)	1.6989	0.19
Weight	-0.00735	0.00380	0.99	(0.99, 1.00)	3.7404	0.05
y = severe anemia						
Distance (total length)	0.000027	0.000024	1.00	(1.00, 1.00)	1.2630	0.26
Age	0.0857	0.0791	1.09	(0.93, 1.27)	1.1739	0.28
Number of previous gestations	-0.4176	0.3189	0.66	(0.35, 1.23)	1.7140	0.19
Weight	-0.0623	0.0273	**0.94**	**(0.89, 0.99)**	5.2070	**0.02**

Table 6: Kulldorff Scan Statistic – High Rates Only

Level	Cluster Number	No. Id's in Cluster	Log Likelihood Ratio	P-value	Observed	Expected	Relative Risk
Any Subclinical Malaria							
Total Study Area	1	5	5.035	0.81	3	0.57	5.42
	2	7	5.035	0.94	3	0.57	5.42
	3	8	5.035	0.94	3	0.57	5.42
Urban	1	182	6.377	0.21	22	13.58	3.89
	2	20	0.184	1.00	2	1.34	1.53
Rural*	1	7	4.911	0.68	3	0.59	5.22
	2	8	4.911	0.95	3	0.59	5.22
	3	186	4.601	0.95	24	14.22	1.98
Subclinical *P. vivax*							
Total Study Area	1	222	6.168	0.55	27	14.88	2.15
Urban	1	29	6.845	0.16	8	2.26	4.74
	2	5	5.840	0.48	3	0.45	7.41
Rural*	1	186	6.478	0.36	23	11.97	2.40
	2	2	3.614	1.00	2	0.33	6.17
	3	3	3.614	1.00	2	0.33	6.17
Subclinical *P. falciparum*							
Total Study Area	1	284	11.048	**0.01**	16	5.15	5.70
Urban	1	41	2.836	0.88	3	0.74	8.67
Rural*	1	284	10.029	**0.01**	16	2.68	6.04
	2	11	2.072	1.00	2	5.42	5.82
	3	15	2.072	1.00	2	5.42	5.82
Low Birth Weight							
Total Study Area	1	2	5.150	0.81	2	0.15	13.34
Urban	1	2	5.243	0.46	2	0.15	14.42
	2	46	4.690	0.86	7	2.13	4.43
	3	7	3.405	0.99	2	0.23	9.58
Rural*	1	2	5.150	0.81	2	0.16	13.45
	2	3	5.150	0.95	2	0.16	13.45
	3	3	5.150	0.95	2	0.16	13.45
Premature							
Total Study Area	1	10	5.301	0.88	3	0.28	11.66
	2	21	4.856	0.91	4	0.61	7.33
Urban	1	10	5.636	0.40	3	0.27	14.44
	2	4	3.671	0.86	2	0.18	13.10
Rural*	1	21	4.898	0.77	4	0.62	7.66
	2	5	4.322	0.92	2	0.14	15.33
	3	58	2.027	1.00	3	0.80	4.14
Anemia							
Total Study Area*	1	636	9.668	**0.046**	144	114.93	1.56
	2	6	4.521	1.00	4	1.30	3.11
	3	6	4.521	1.00	4	1.30	3.11

Level	Cluster Number	No. Id's in Cluster	Log Likelihood Ratio	P-value	Observed	Expected	Relative Risk
Anemia							
Urban*	1	12	7.591	0.09	7	2.43	3.03
	2	7	5.387	0.70	5	1.73	2.99
	3	5	4.295	0.97	4	1.39	2.97
Rural*	1	21	4.898	0.77	4	0.62	7.66
	2	5	4.322	0.92	2	0.14	15.33
	3	58	2.027	1.00	3	0.80	4.14

*More than 3 clusters were calculated; the 3 reported had the smallest p-values

Table 7: Loss of information by numbers

Data (variables)	N observations	N observation with houses	N observations with houses with routes
Houses (Spatially located)	1,554	--	--
Houses with Connecting Routes	1,548	--	--
PCR (all subclinical malaria outcomes)	1,340	746	744
Delivery (place of delivery, low birth weight)	1,032	816	814
Newborn (prematurity)	1,011	802	800
Anemia (anemia, severe anemia, hemoglobin)	1,025	812	810

Appendix B: Figures

Figure 1: Q-Q plot for any subclinical malaria (left- no, right- yes)

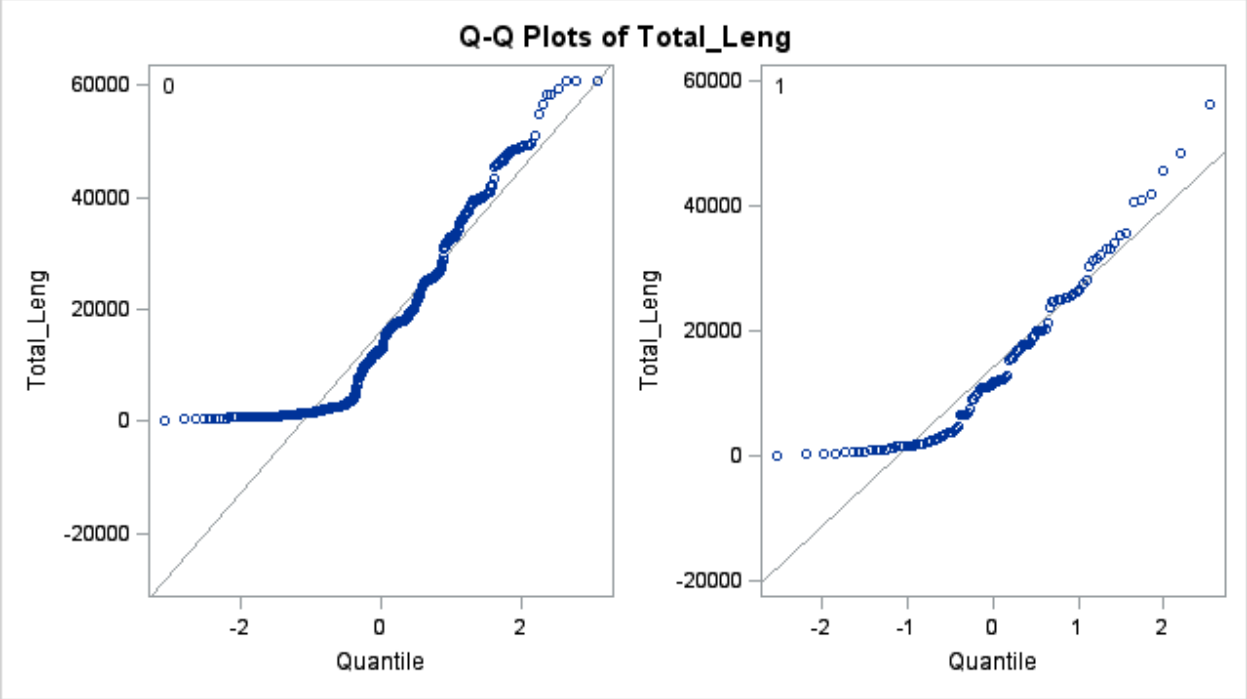


Figure 2: K function for positive malaria cases (any species) (in meters)

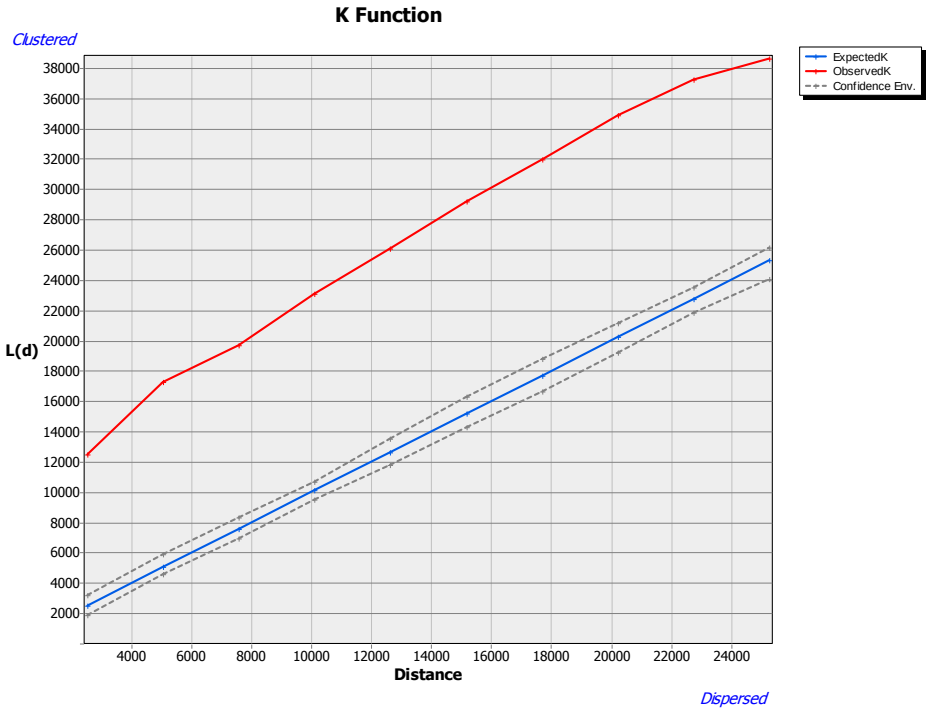


Figure 3: K function for positive *P. vivax* cases (in meters)

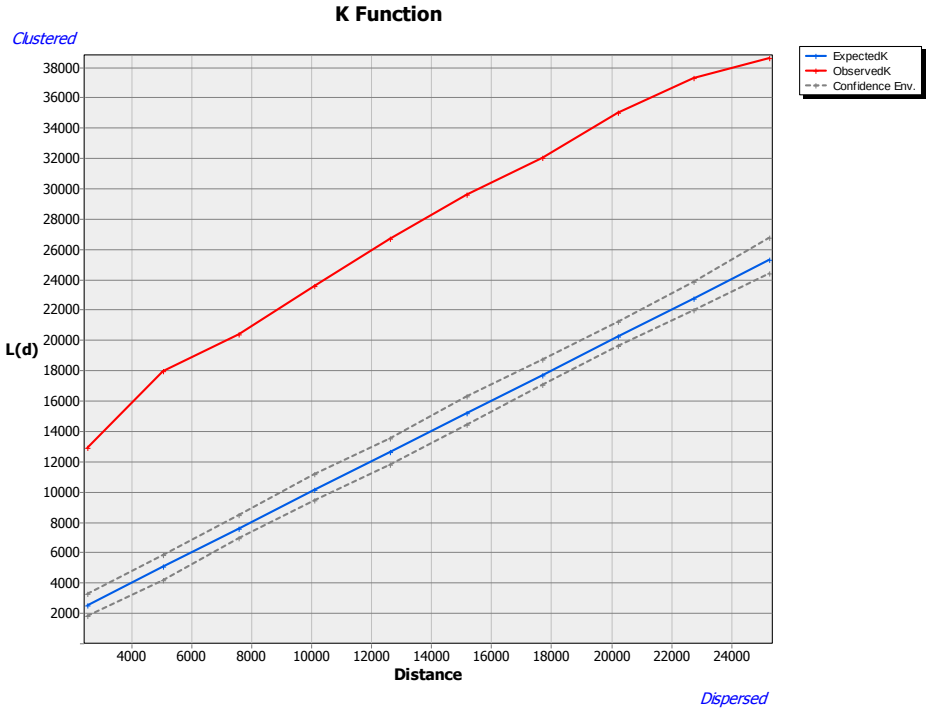


Figure 4: K function for positive *P. falciparum* cases (in meters)

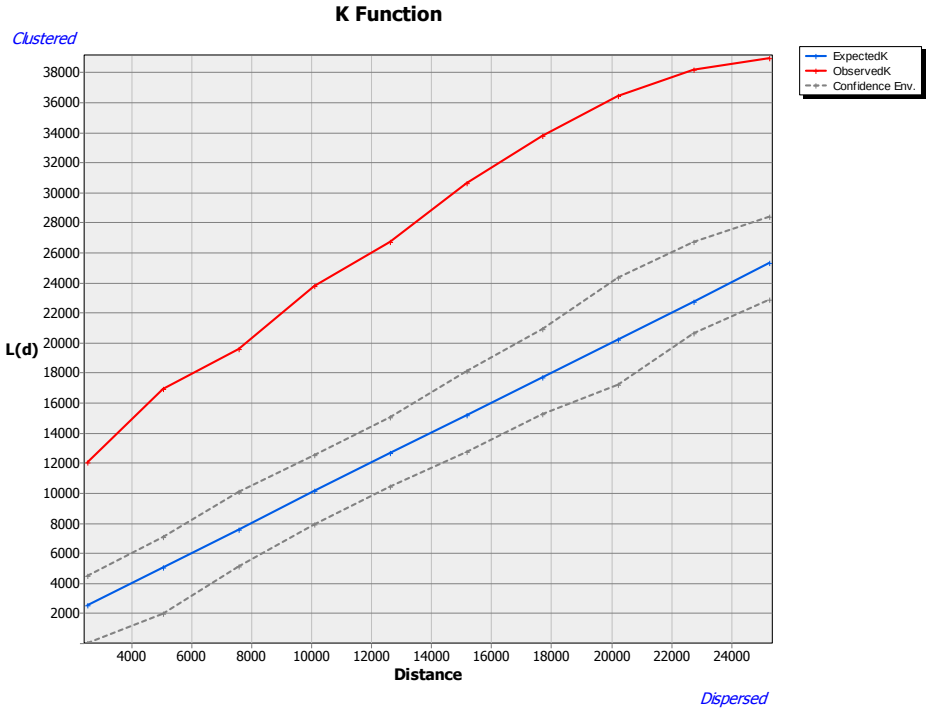


Figure 5: K function for low birth weight (in meters)

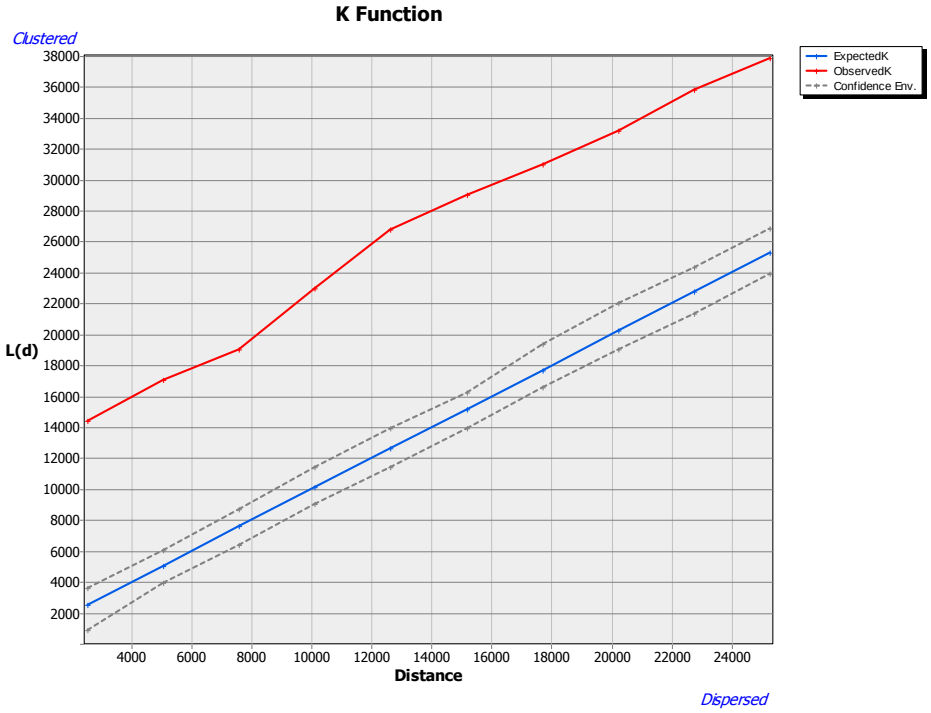


Figure 6: K function for prematurity (in meters)

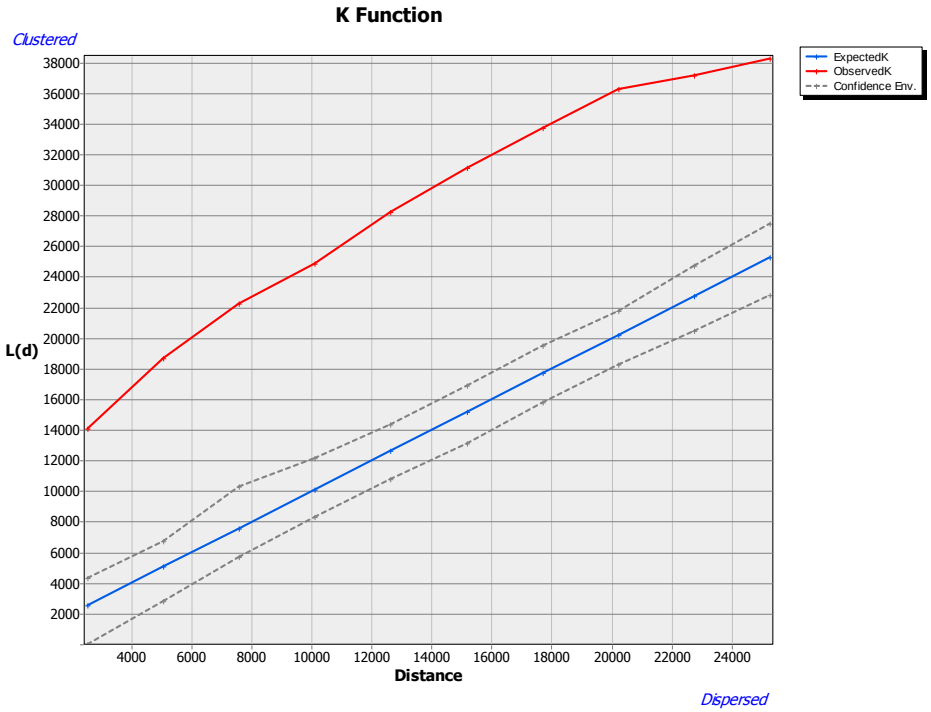
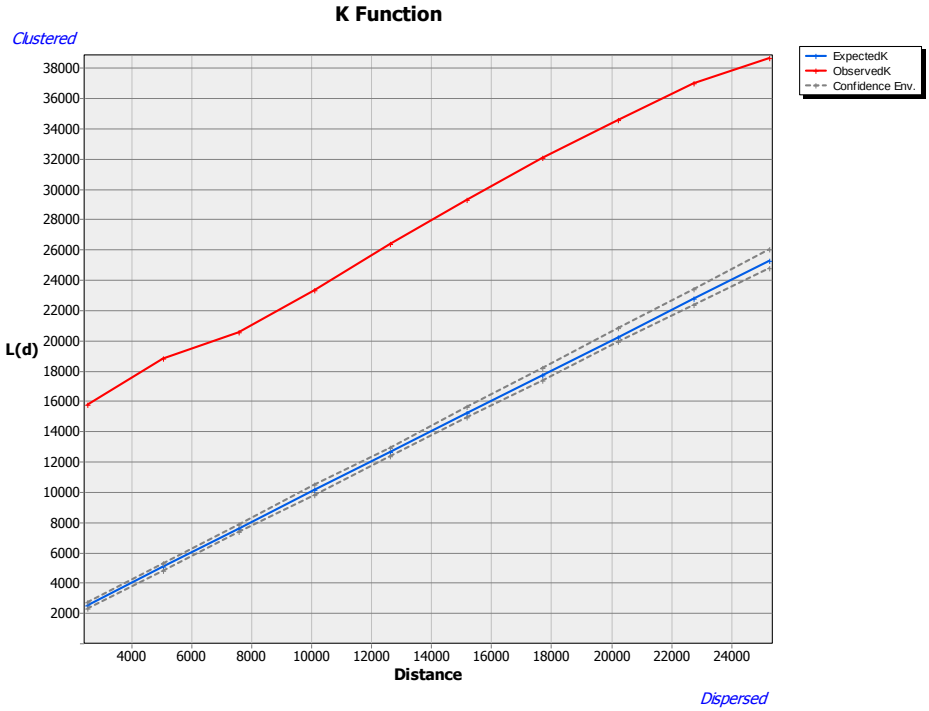
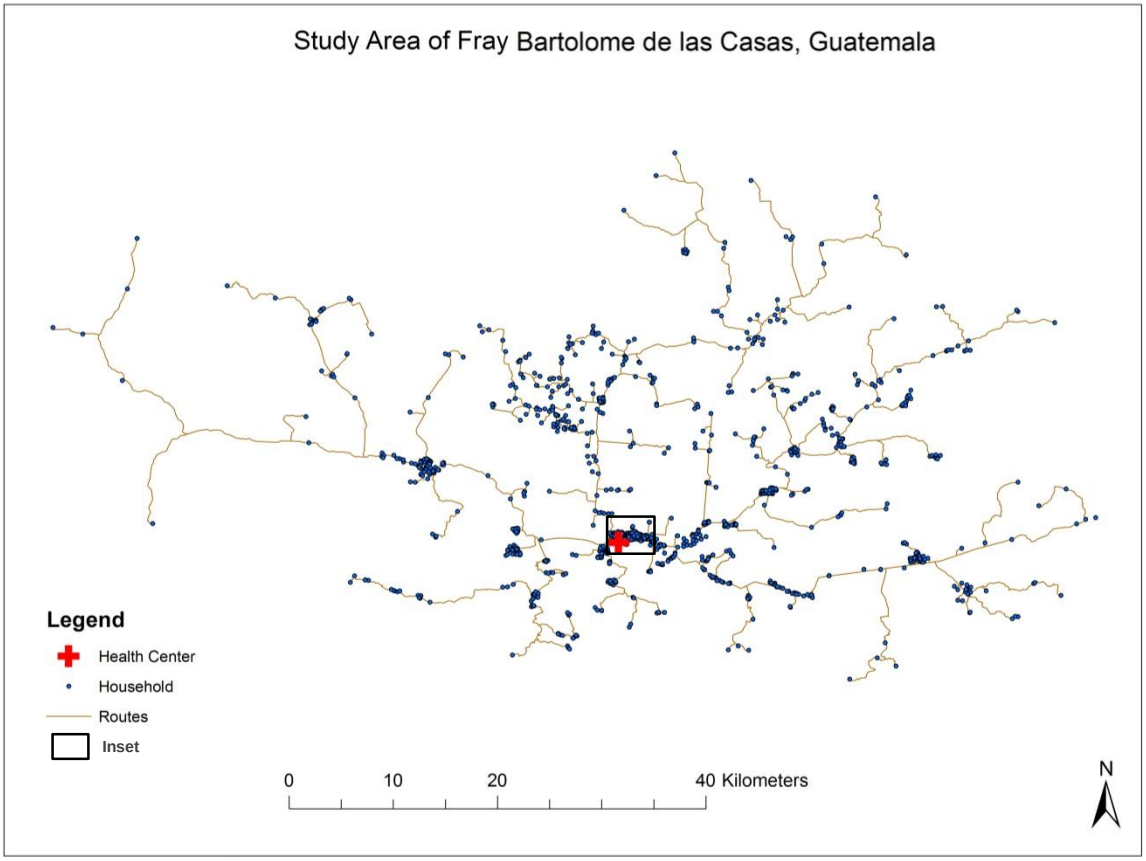


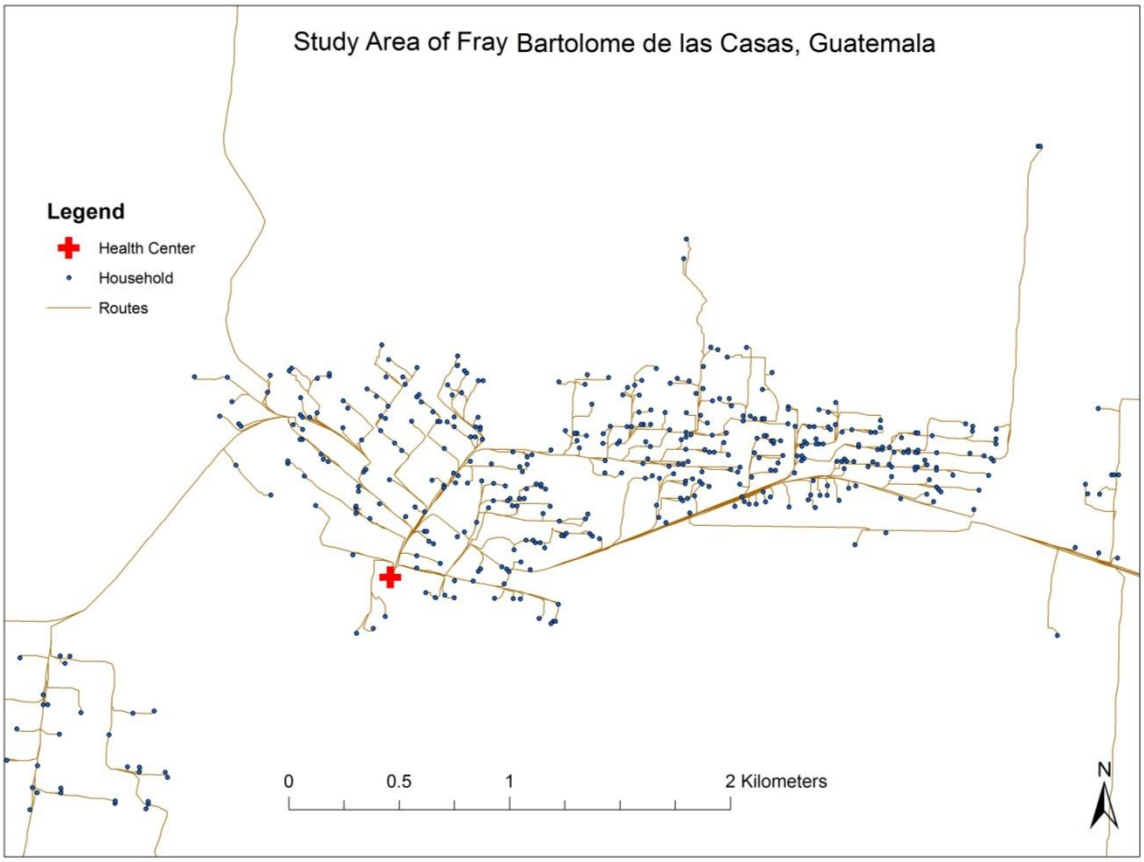
Figure 7: K function for anemia (in meters)



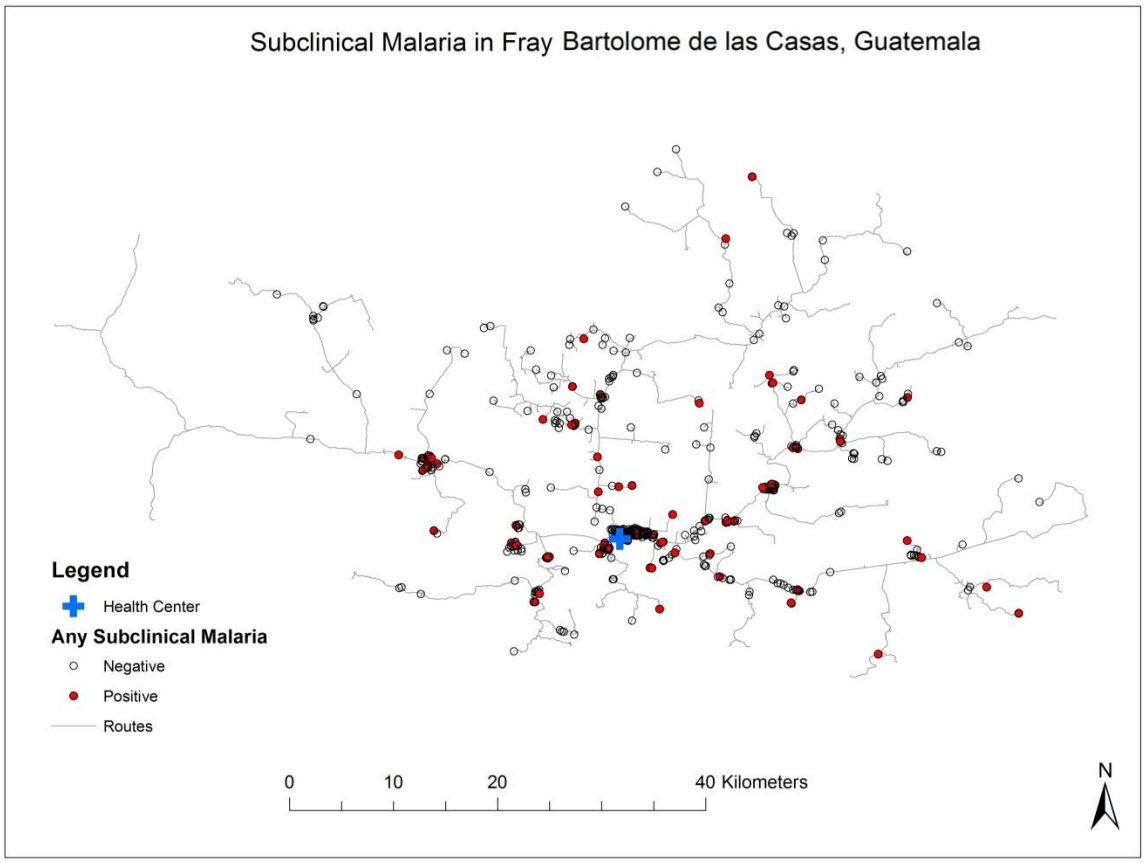
Appendix B: Maps

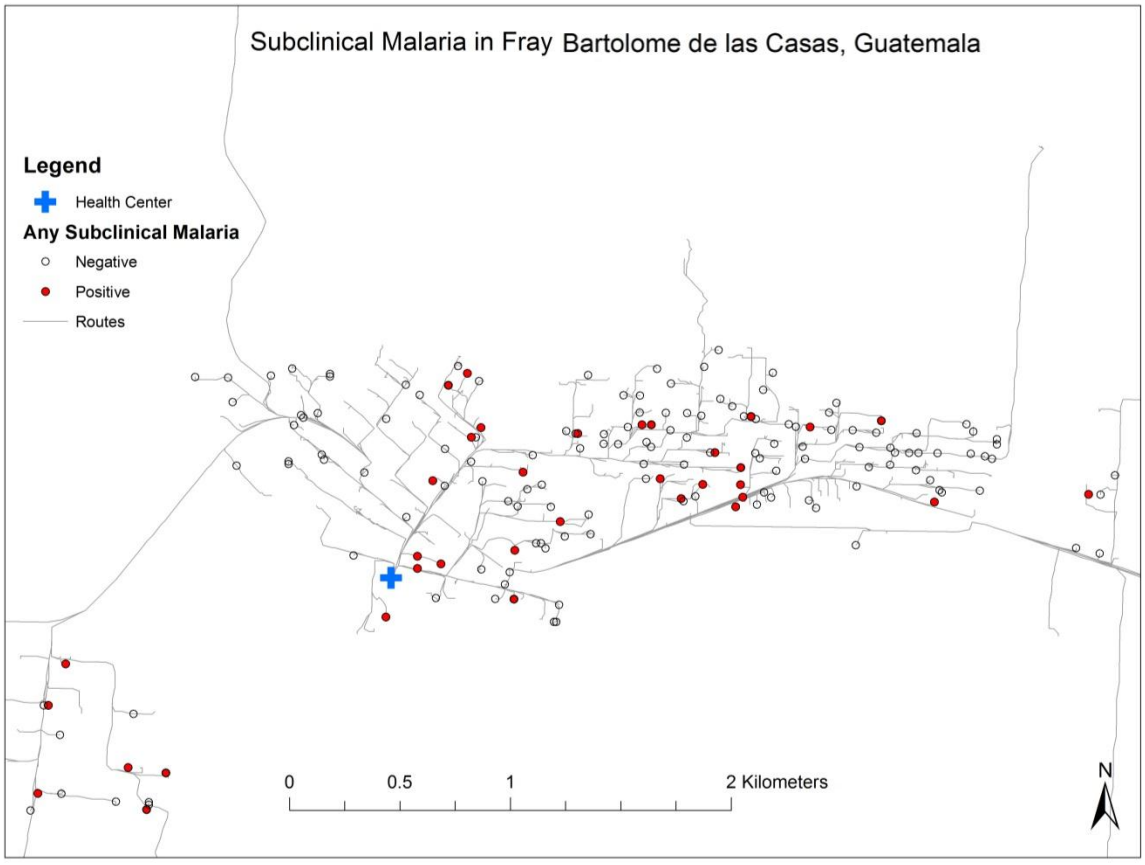
Map 1: Base map of study area (top) and urban center (bottom)



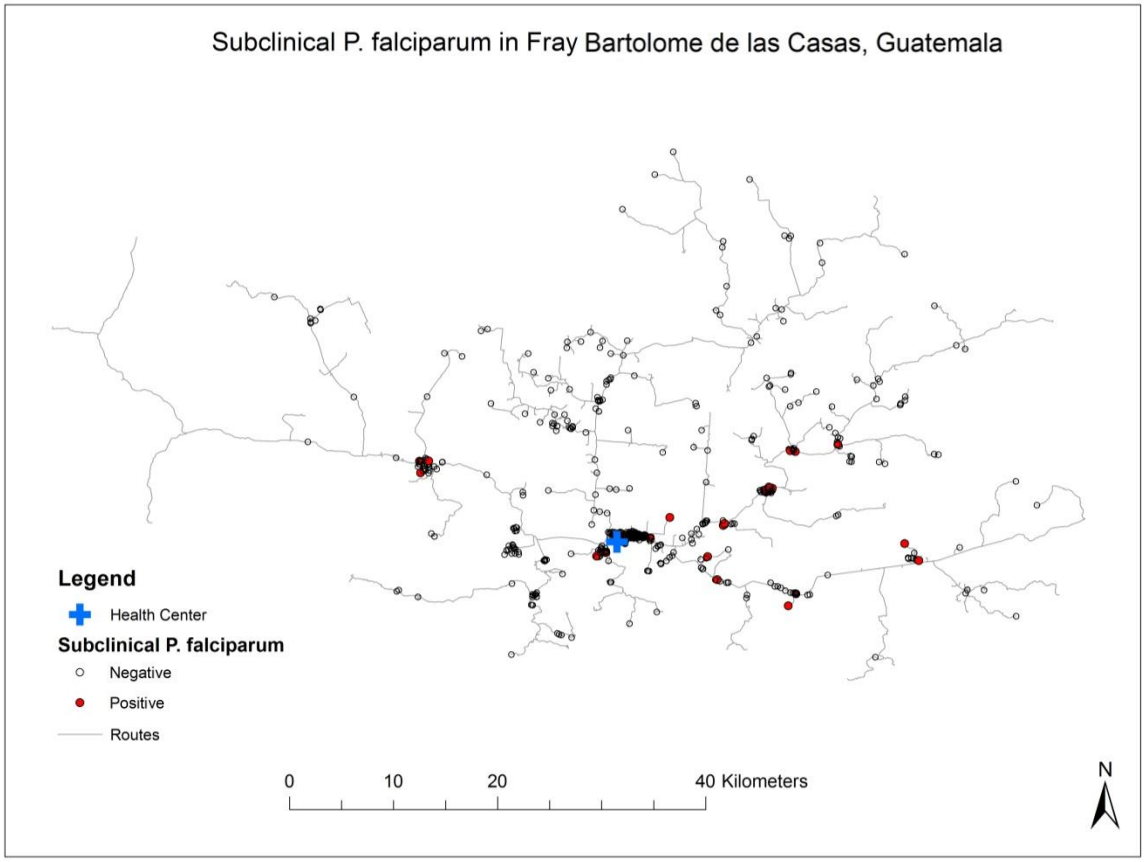


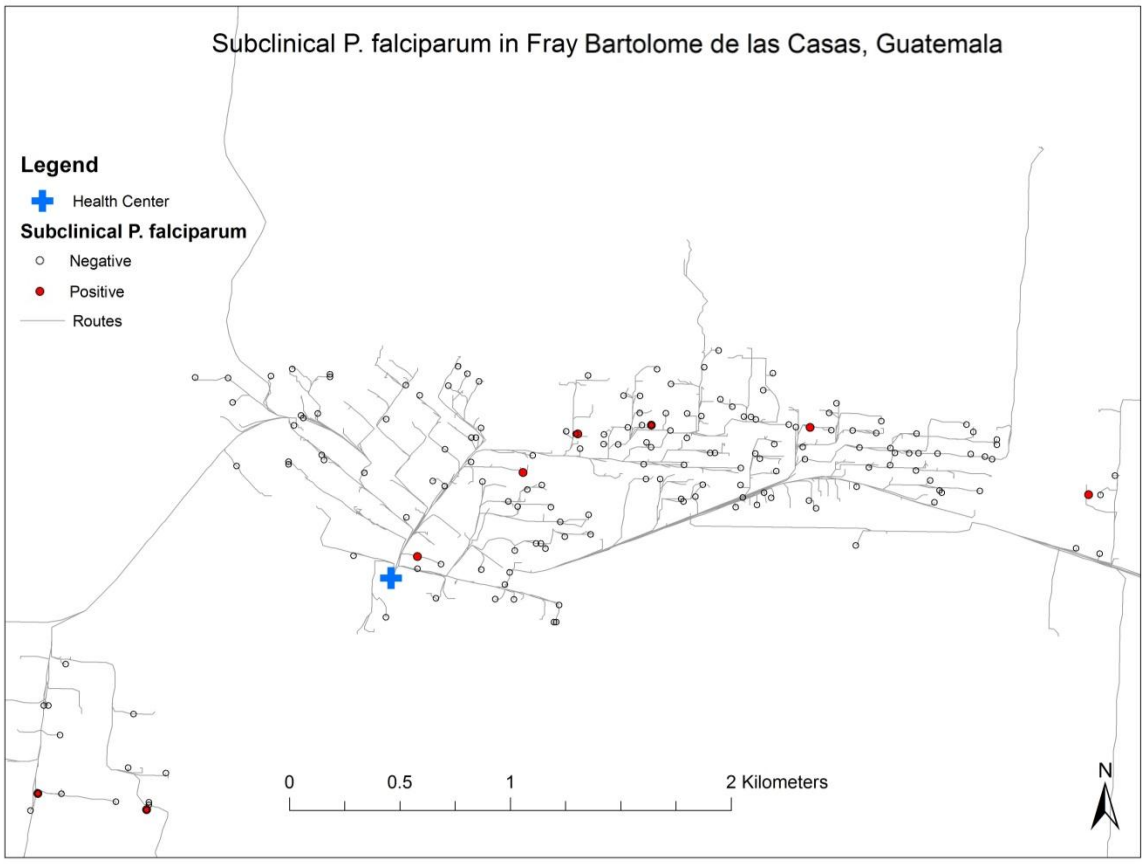
Map 2: Any subclinical malaria in study area (top) and urban center (bottom)



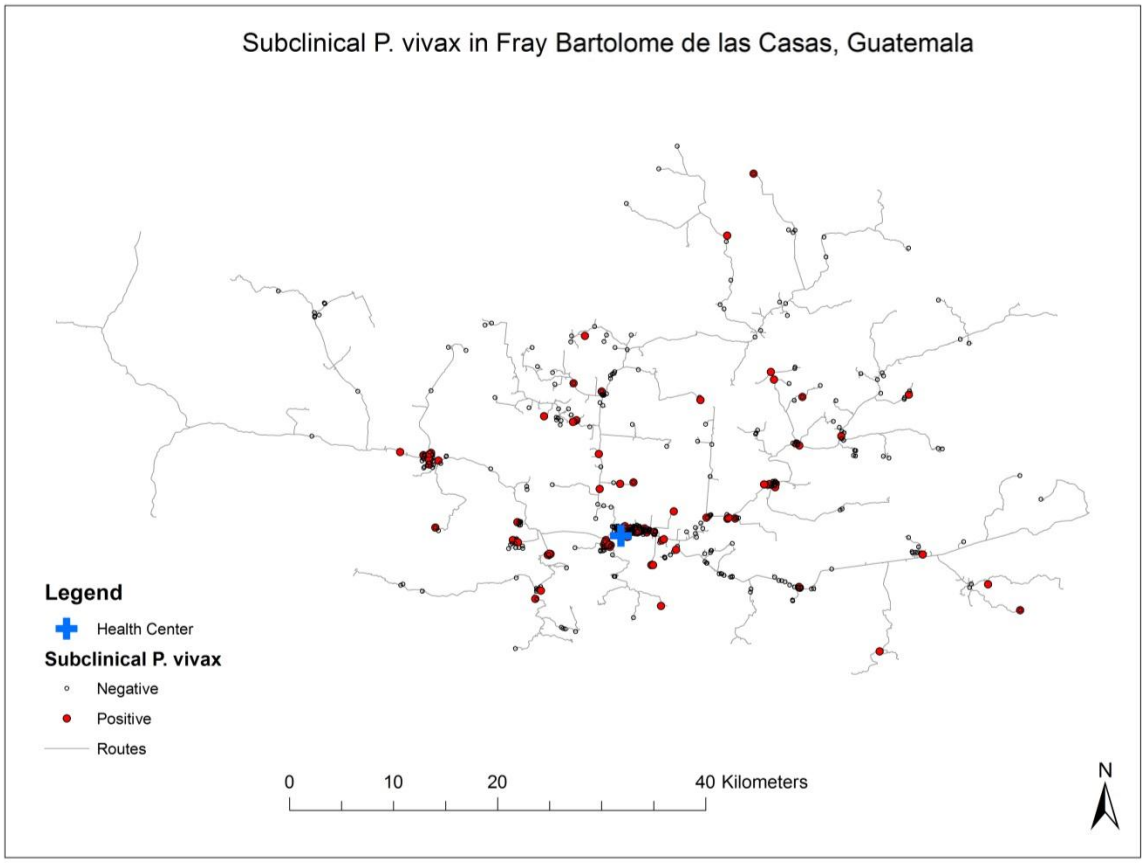


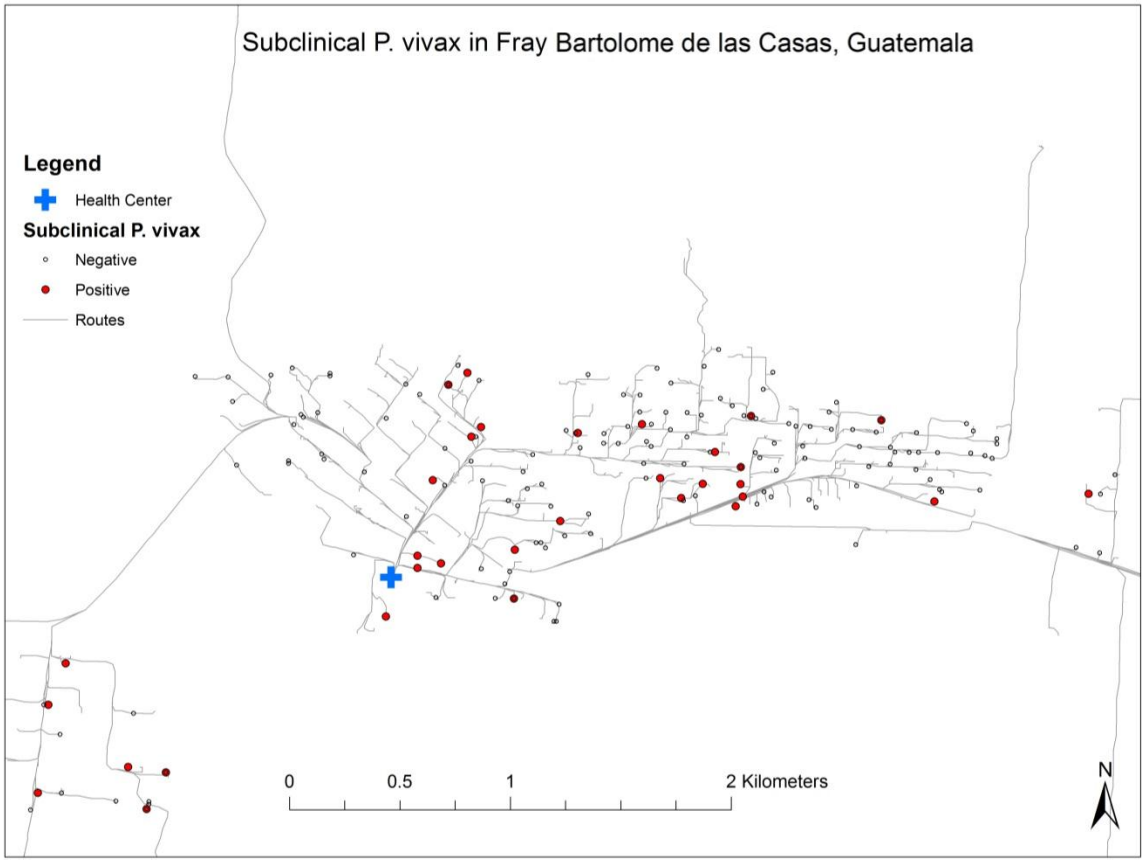
Map 3: Subclinical *P. falciparum* in study area (top) and urban center (bottom)



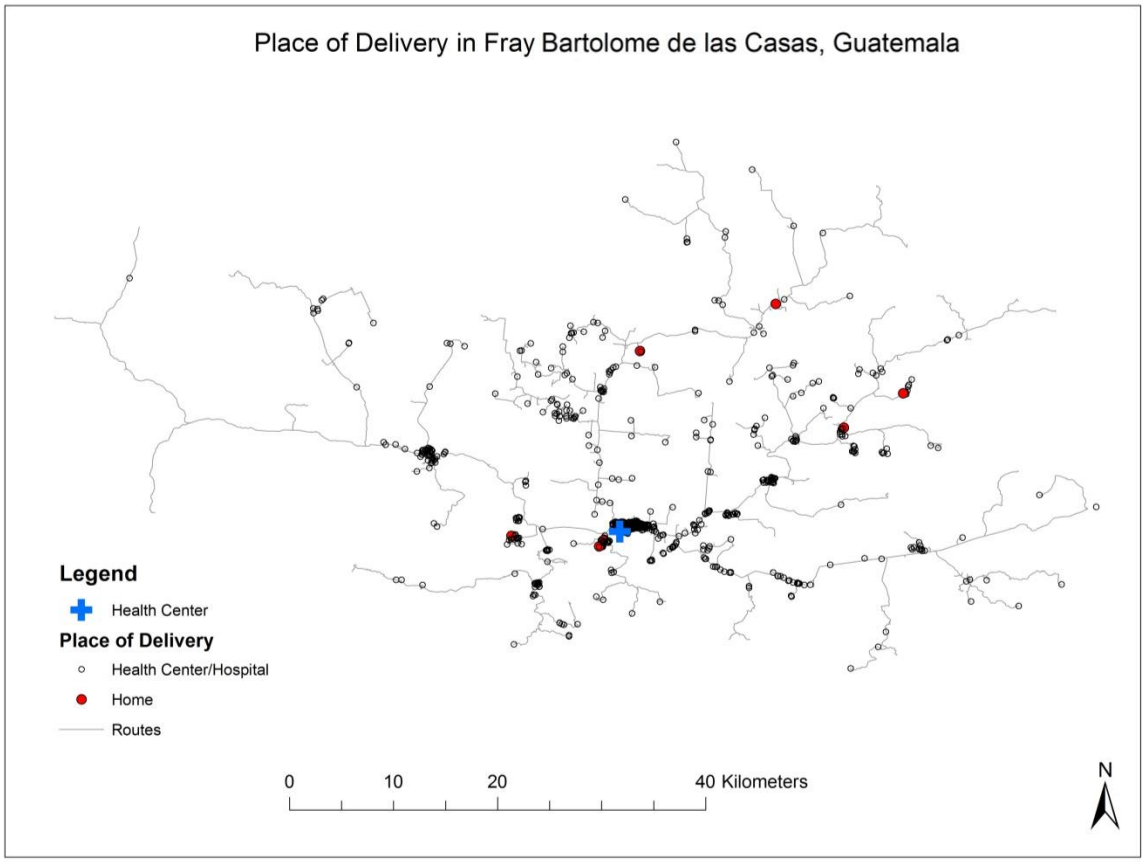


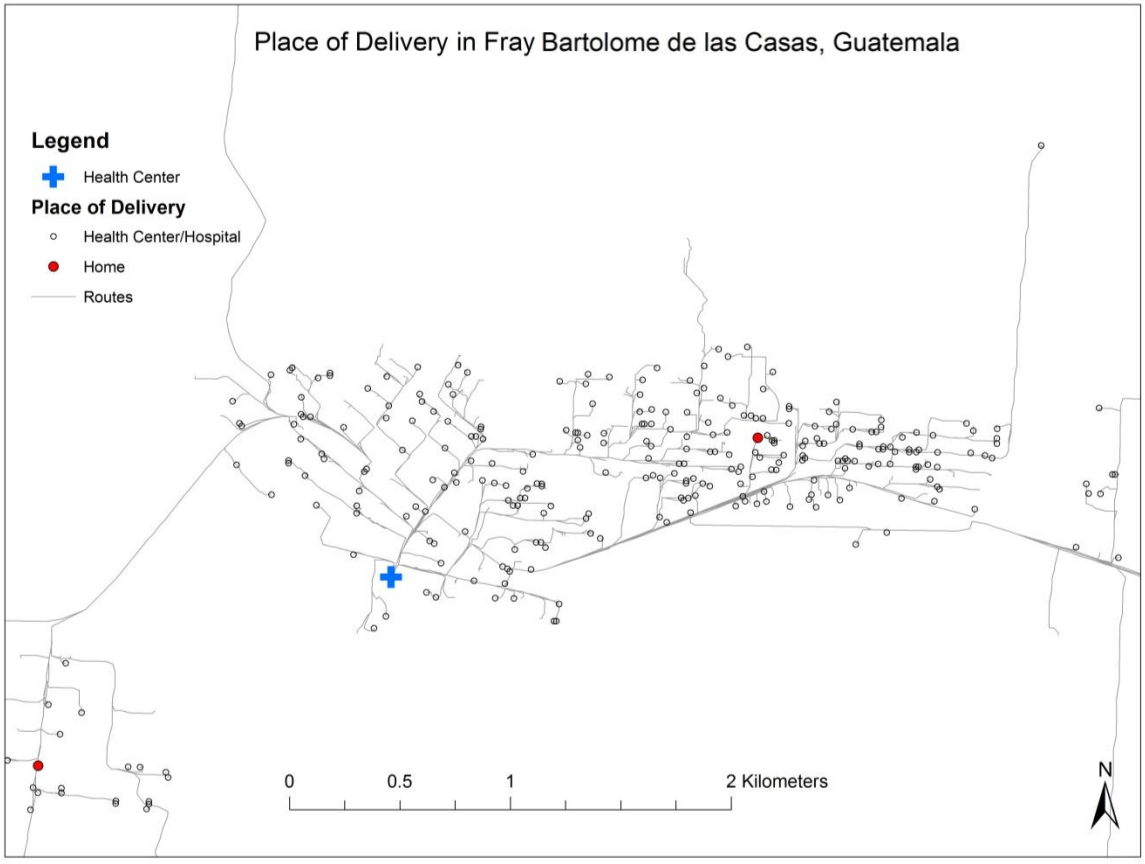
Map 4: Subclinical *P. vivax* in study area (top) and urban center (bottom)



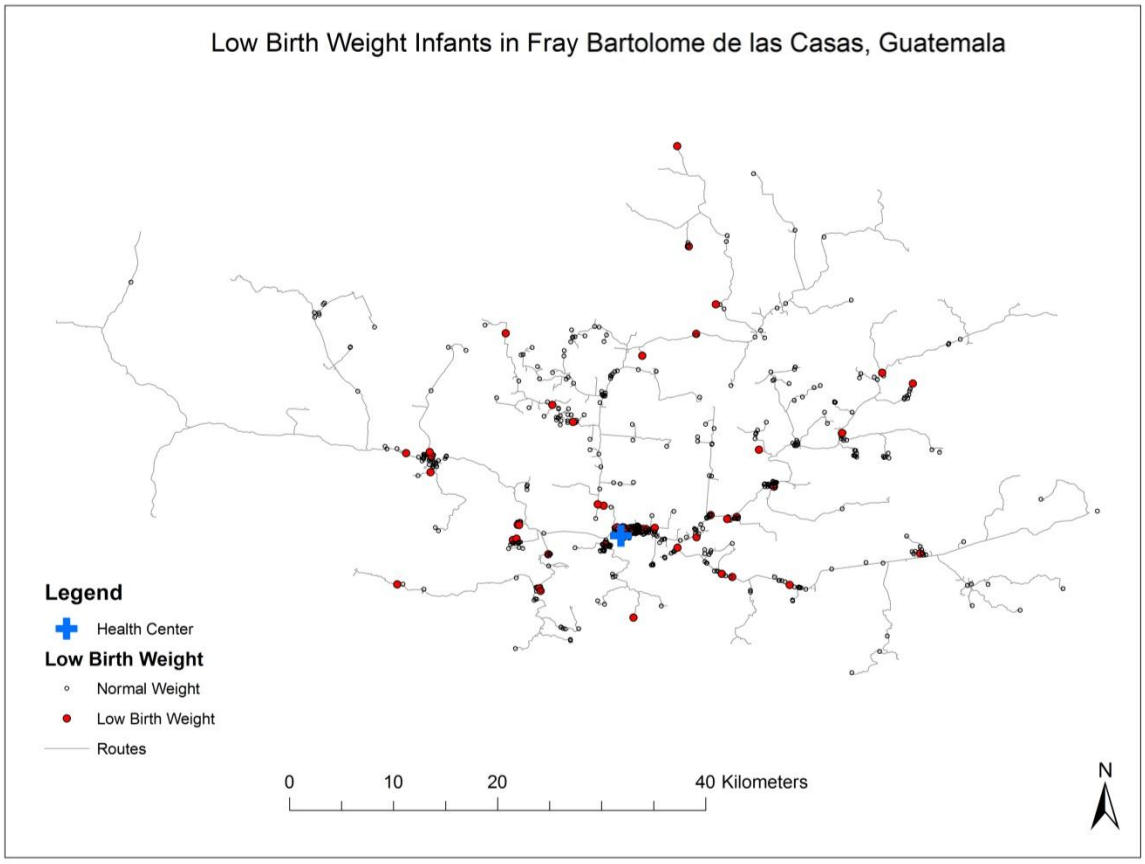


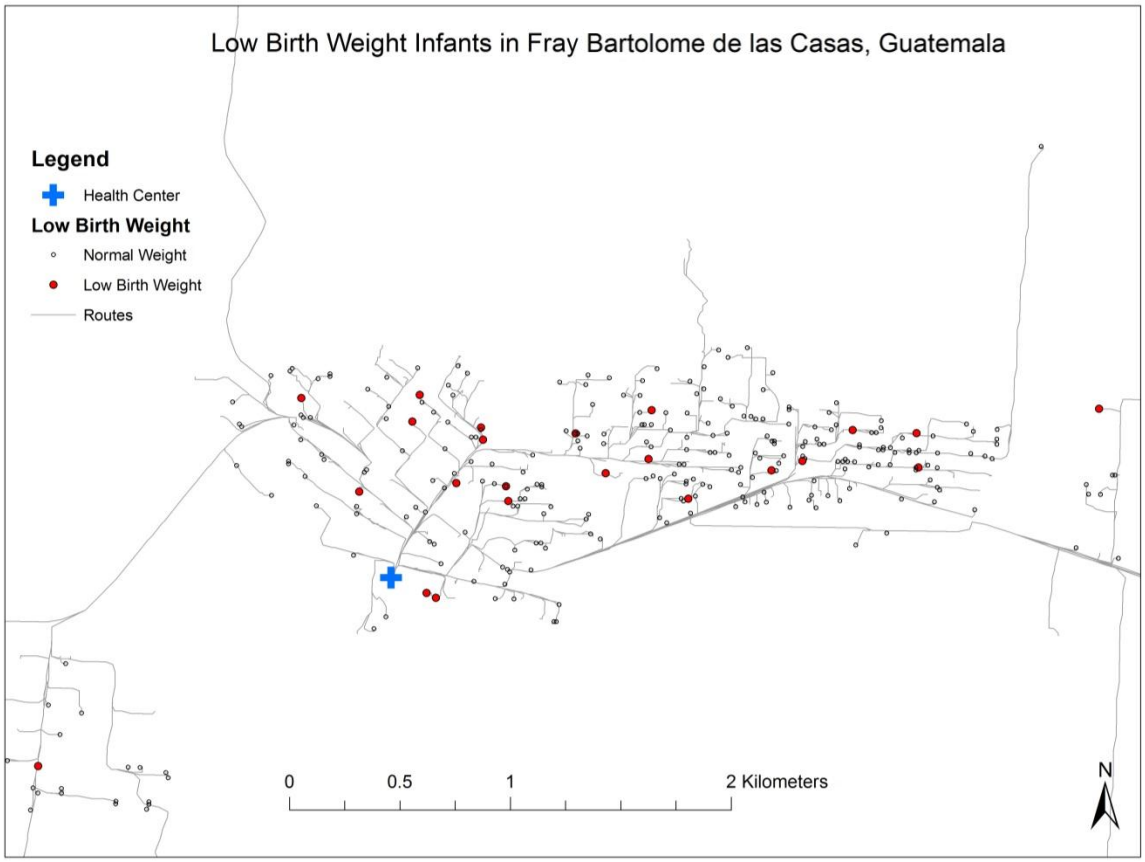
Map 5: Place of delivery in study area (top) and urban center (bottom)



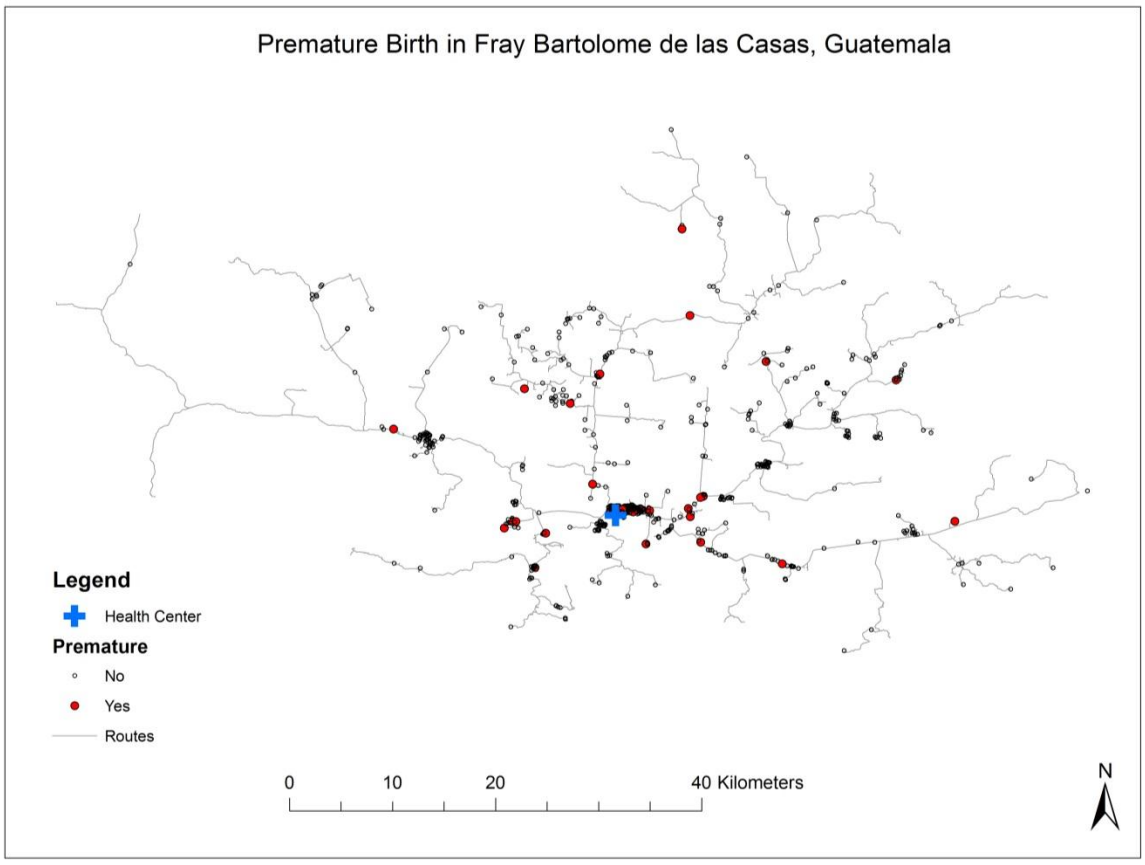


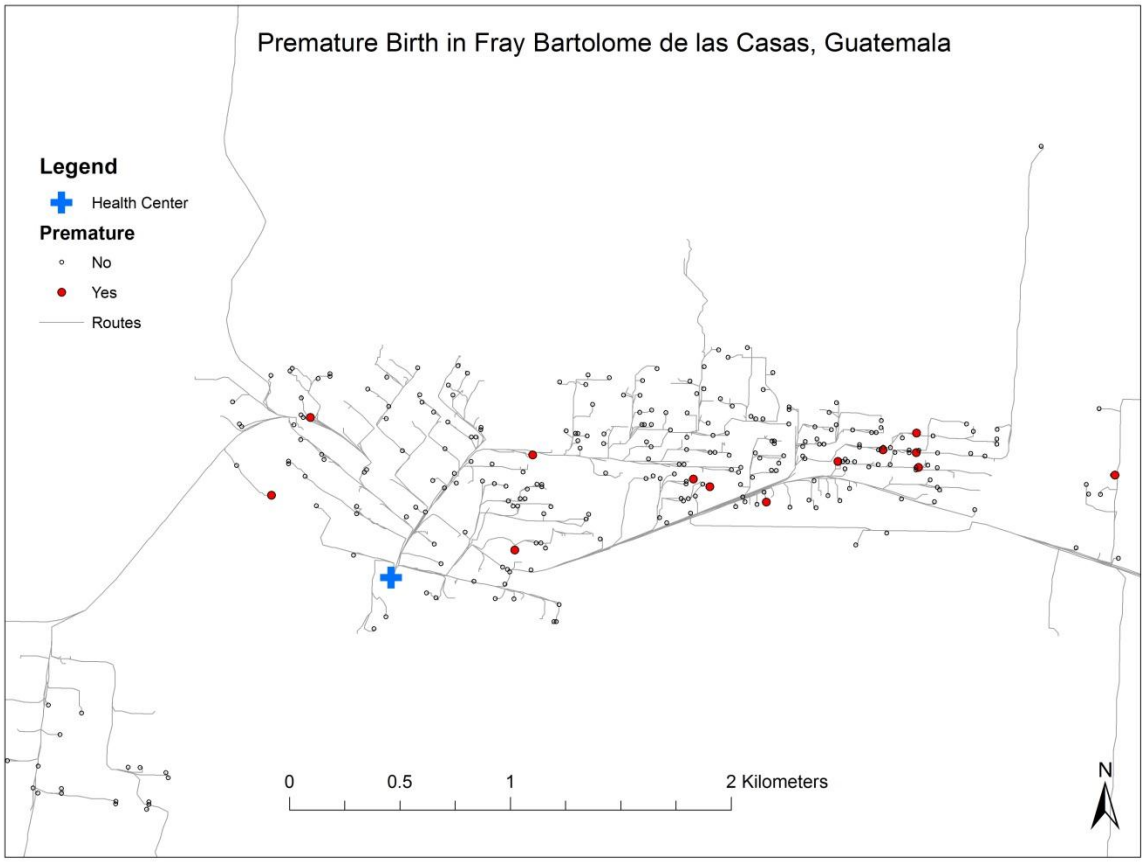
Map 6: Low Birth Weight in study area (top) and urban center (bottom)



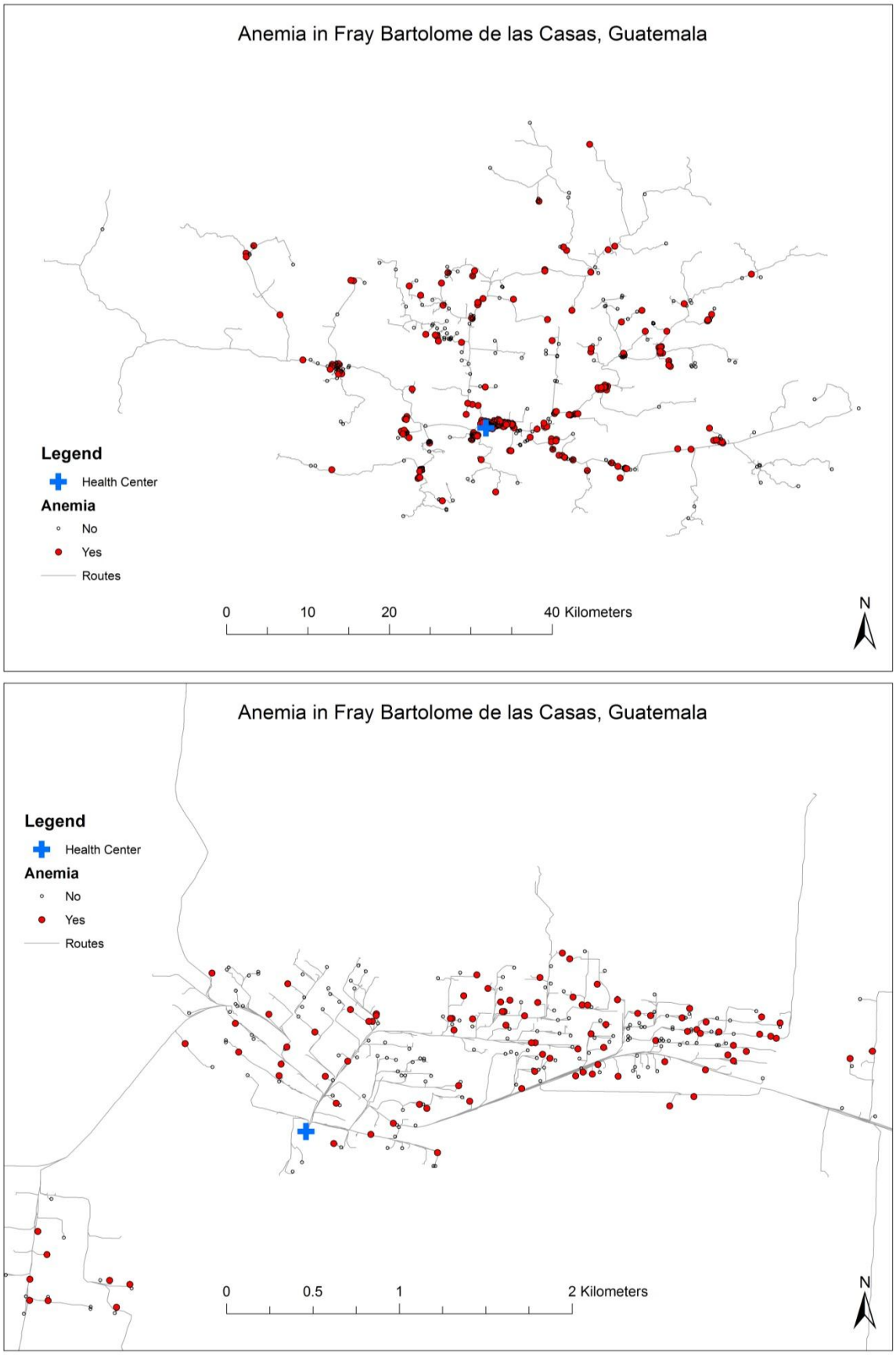


Map 7: Premature birth in study area (top) and urban center (bottom)

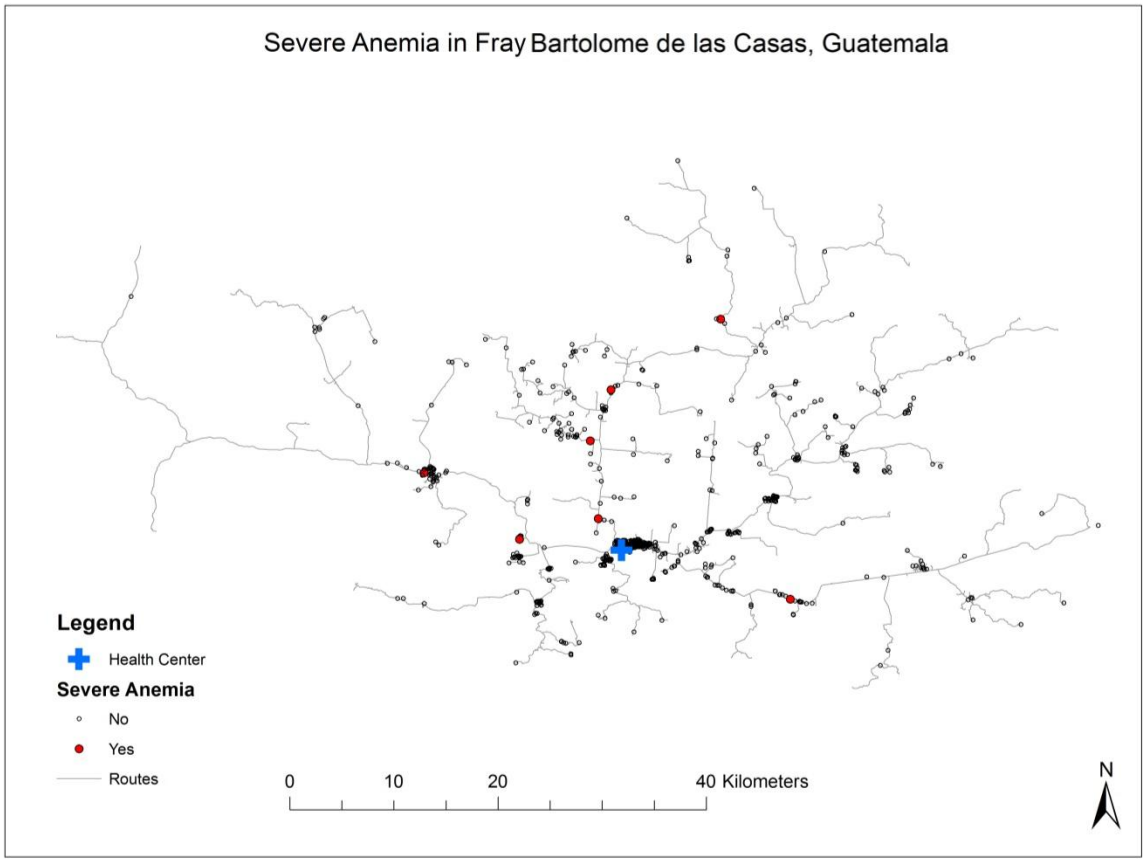


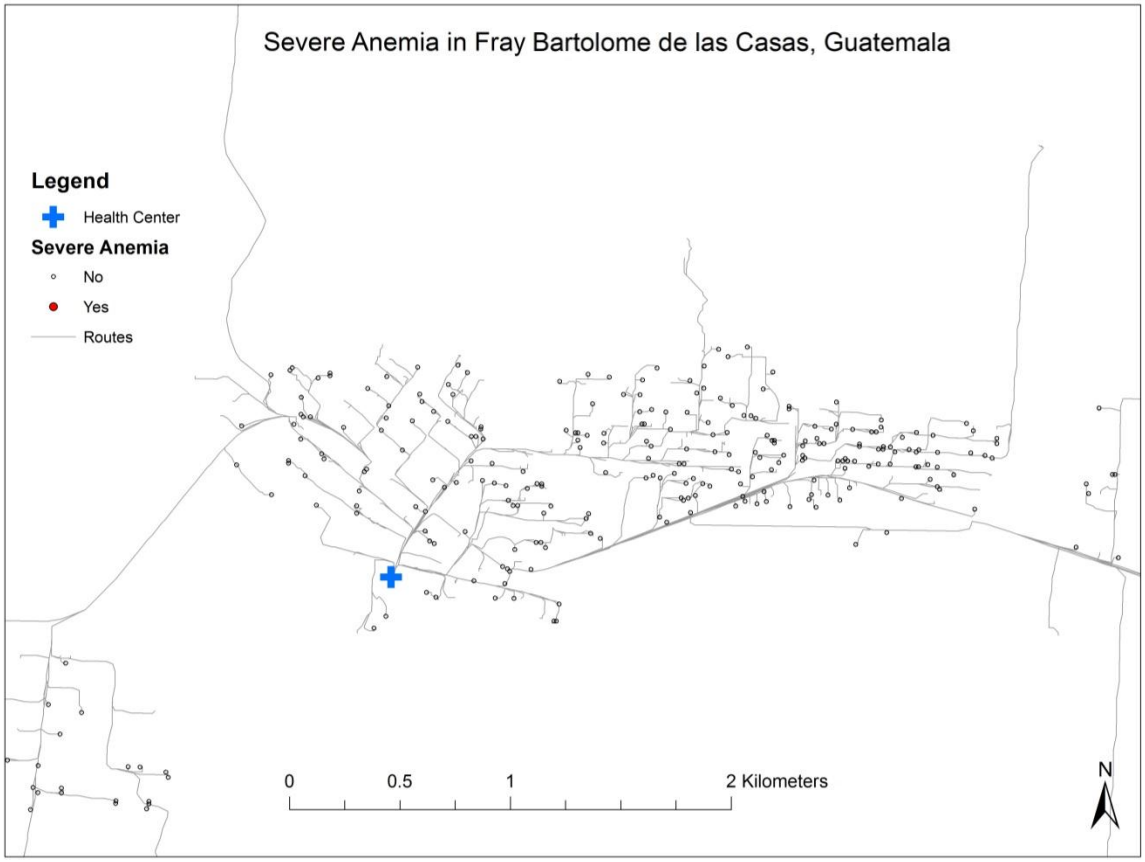


Map 8: Anemia in study area (top) and urban center (bottom)

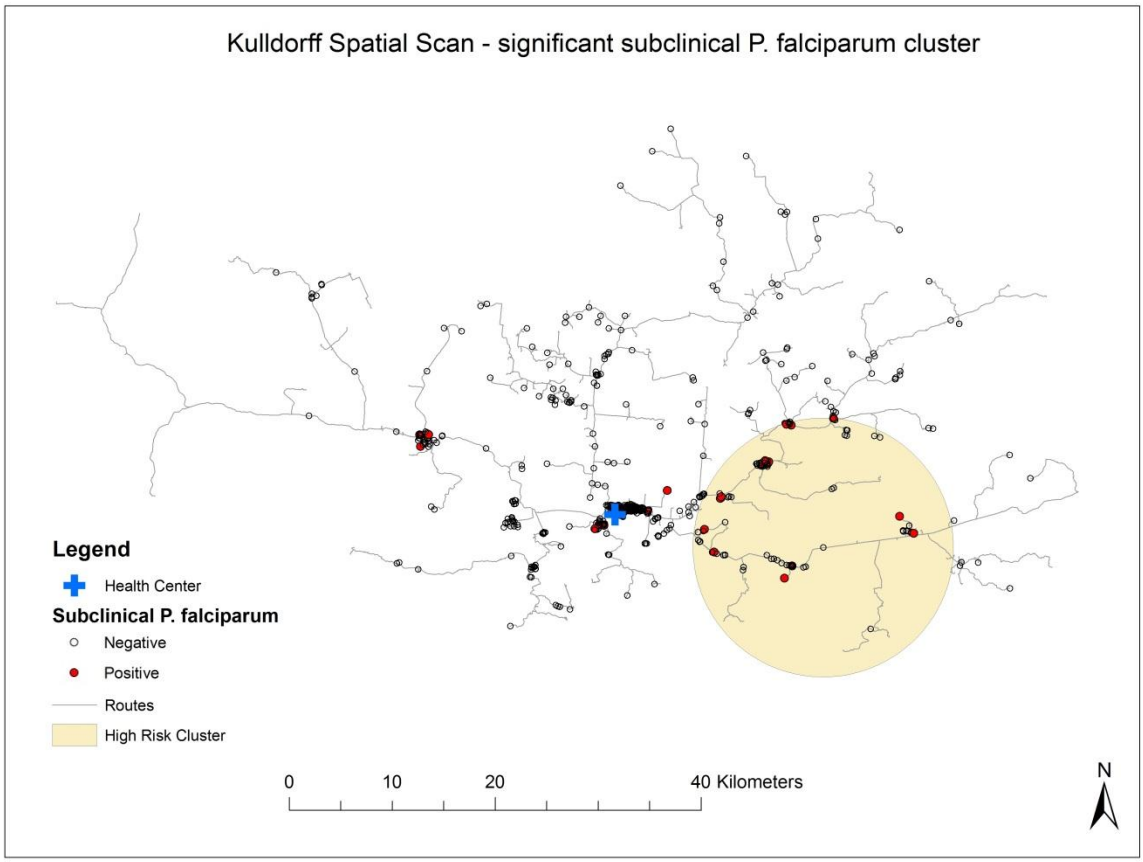


Map 9: Severe anemia in study area (top) and urban center (bottom)

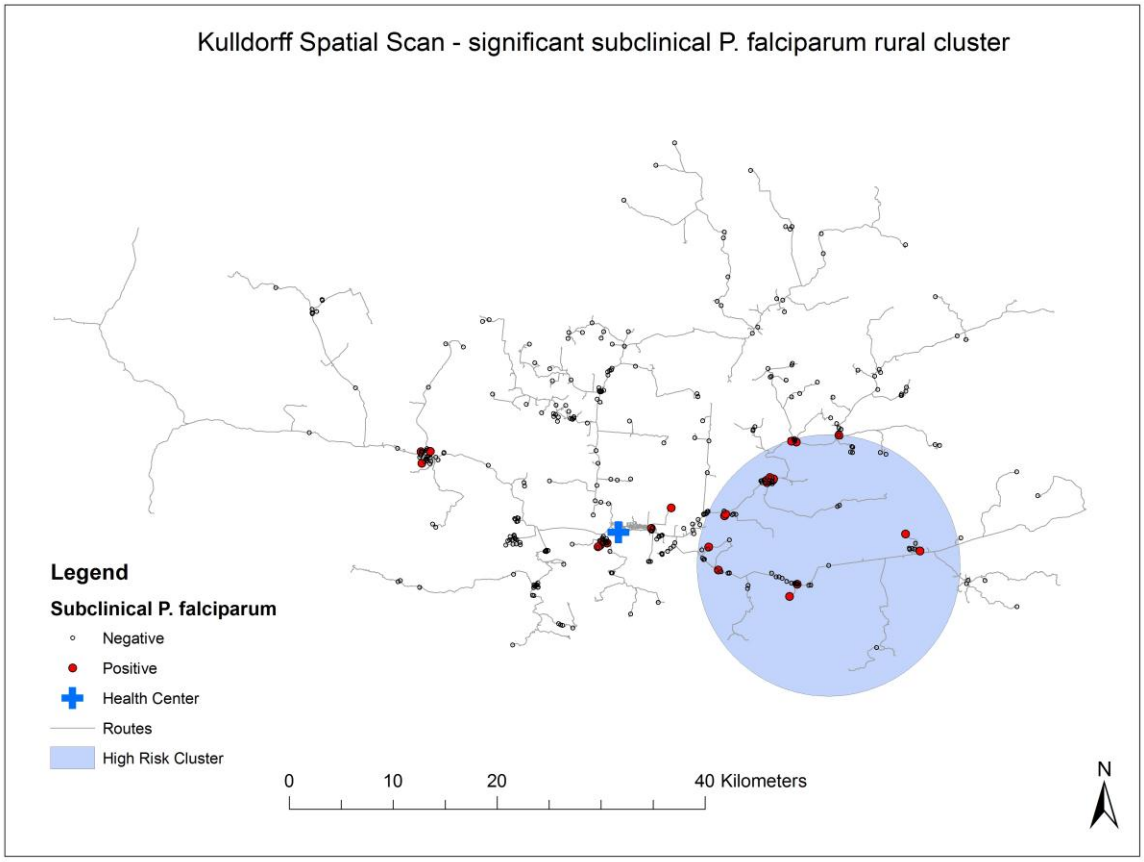




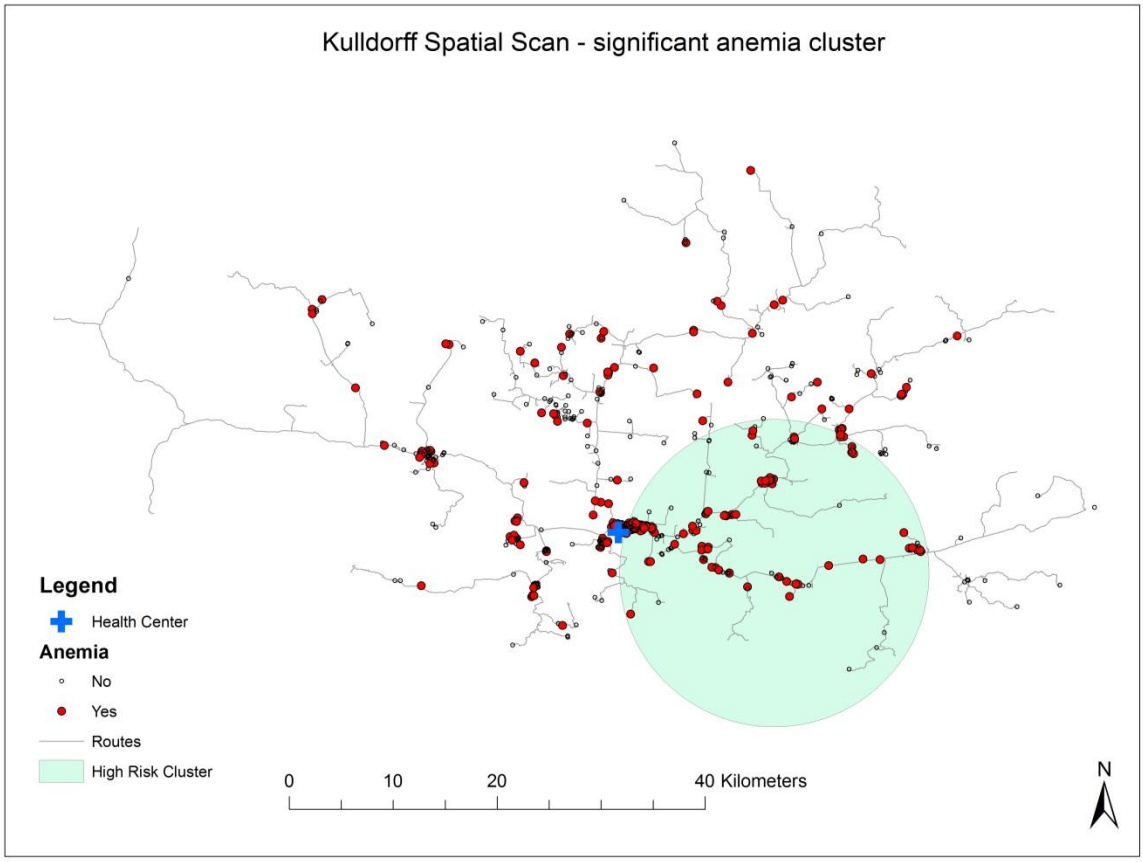
Map 10: Subclinical *P. falciparum* cluster using Kulldorff spatial scan statistic, total area



Map 11: Subclinical *P. falciparum* cluster using Kulldorff spatial scan statistic, rural only



Map 12: Anemia cluster using Kulldorff spatial scan statistic



Map 13: Ripley's K Function study area boundary