USE OF GIS FOR ANALYSIS OF COMMUNITY HEALTH WORKER PATIENT REGISTRIES FROM CHONGWE DISTRICT, A RURAL LOW-RESOURCE SETTING, IN LUSAKA PROVINCE, ZAMBIA

by

Mine Metitiri, MPH, CPH

A Thesis Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

December 2014

TABLE OF CONTENTS

DEDICATION

This is dedicated to the Community Health Workers and Volunteers serving on the frontlines of healthcare delivery in Zambia and developing regions globally.

"Everybody can be great, because anybody can serve. You don't have to have a college degree to serve. You don't have to make your subject and verb agree to serve. You only need a heart full of grace. A soul generated by love."
- Dr. Martin Luther King Jr.

ACKNOWLEDGEMENTS

I would like to acknowledge my committee members Dr. Darren Ruddell and Dr. Daniel Warshawsky. Many thanks to my mentors Dr. Travis Longcore for his support and guidance on this project and Dr. Pascalina Chanda-Kapata for her guidance also and whose work in Chongwe district served as the inspiration for this thesis.

Thank you to my family and friends for their constant support and encouragement throughout this process, without whom I would not have gotten this far, most especially William Ngosa, Jessie Pinchoff, and Colin Mattison for sharing their data analysis and GIS expertise when answering my questions and Dr. Dela Kusi-Appouh for reading all of my drafts. I am forever grateful.

Thank you also to the Zambian Ministry of Health staff at all government levels, the staff at the Central Statistical Office in Lusaka, the Surveyor General's Office in Lusaka, and PATH's Lusaka offices for sharing their datasets and providing information.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACT	Artemisinin-base Combination Therapy
CD	Communicable Diseases
CHW	Community Health Worker
CSO	Central Statistical Office
DHIS$_2$	District Health Information System version 2
eHealth	Electronic Health
GDP	Gross Domestic Product
GIS	Geographic Information System
GIST	Geographic Information Systems and Technology
ICCM	Integrated Childhood Case Management
ICT	Information and Communication Technologies
IRS	Indoors Residual Spray
ITNs	Insecticide Treated Nets
LLINs	Long-lasting Insecticide Nets
LMICs	Low-Middle Income Countries
LR	Likelihood Ration
mHealth	Mobile Health
MOH	Zambia Ministry of Health
NCD	Non-communicable Diseases
NGO	Non-Governmental Organization
NMCP	National Malaria Control Program
ORS	Oral Rehydration Solution
PDA	Personal Digital Assistant
RDT	Rapid Diagnostic Test

RHC	Rural Health Center
RS	Remote Sensing
SEA	Standard Enumeration Area
UN	United Nations
UNICEF	United Nations Children's Fund
WHO	World Health Organization

ABSTRACT

The growing accessibility of mobile phones in developing countries has led to increased innovation and utilization of handheld technology in managing health outcomes. Mobile health (mHealth) technologies enabled significant gains in localized data collection methods and increased timeliness in disease surveillance and control programs. Mobile technology has become an important tool for point of care productivity and effective task shifting for Community Health Workers (CHWs) in many developing countries. Concurrently, GIS technology has increasingly been utilized in public health research, planning, monitoring, and surveillance within many developing countries and low-resource settings. This has resulted in opportunities for better understanding of spatial variation of diseases and the correlations with environmental factors.

To better understand community needs and burden of illnesses managed by CHWs, a geospatial analysis at the sub-district level was performed on CHW catchment area health data registries. Risk assessments and cluster analyses were conducted to identify high incidences of fever related illnesses for malaria, diarrhea, and pneumonia in community areas within the rural district area of Chongwe, Zambia. Seventy CHWs recorded 7,674 cases over a time-period of ten months, of which 3,130 cases were geocoded for geospatial analyses. One hundred forty-one village areas within 15 rural health center catchment areas were geocoded and mapped. Results were used to create thematic maps illustrating disease distribution and risks for malaria, pneumonia, and diarrheal illnesses for each sub-district village area manage by CHWs. The use of mobile technology integrated with GIS to manage community health data and the application of GIS to analyze community level data may provide further insight into local area disease distribution, variability, and community needs than systems focused solely on district level data analysis and lacking GIS integration.

CHAPTER ONE: INTRODUCTION

In Zambia, like in many other developing countries, health care workers are in short supply. Community Health Workers (CHW) compensate for shortages in health care workers and serve as the backbone of health care service delivery (Braun et al. 2013; Ministry of Health NHSP 2011-2015). Zambia is a sub-Saharan low-income country located in the Southern hemisphere of the African continent Figure 1.1.

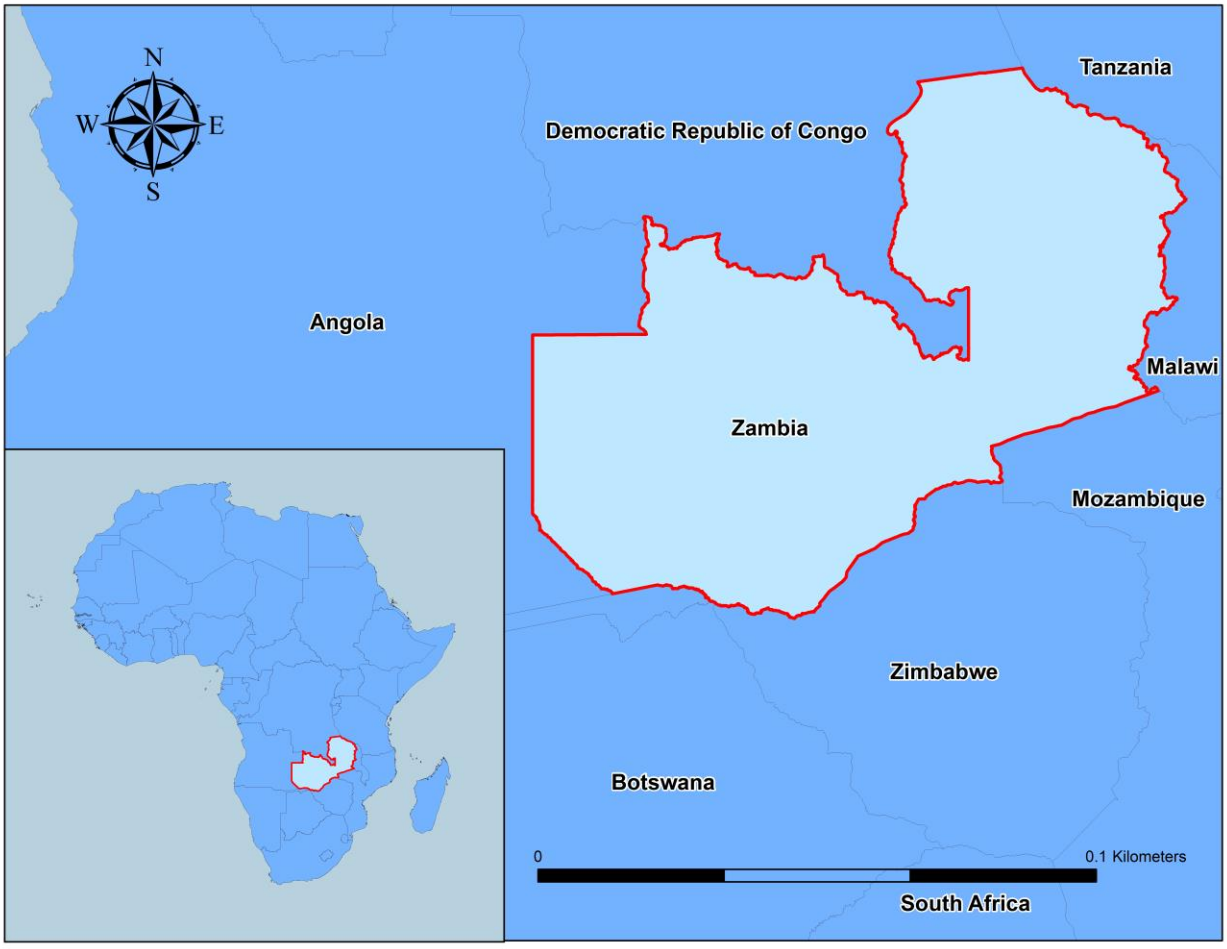


Figure 1.1 Zambia is a sub-Saharan country located in the Southern Hemisphere of Africa

Advances in information and communication technologies (ICT) have provided opportunities for improving community health services delivery both in Zambia and across developing regions. The introduction of ICT devices such as computers and mobile phones, and

training of CHW to use such technologies, has strengthened CHW capacity. It has also opened up opportunities for improving the quality of health service delivery at a community level. Zambian CHW responsibilities include community health data management and reporting of disease burdens to local health centers. The introduction of mobile phones as a tool for community-based case management has led to timely reporting to health facilities and patient feedback (Kamanga et al. 2010, Blaya et al. 2010). The use of mobile devices has been most beneficial for CHW working in rural areas and within communities located far from the district health centers. ICTs are an example of how technology can be used to improve service delivery in rural parts of Zambia.

In Chongwe district, CHWs were provided ICT devices during a pilot program. The aims of the program were to determine the feasibility of usage in a rural setting and to evaluate the effects of mobile devices on CHW activity and retention, health services, and health status of the community. Information collected by each CHW was aggregated and reported on a weekly basis to district health centers. In this reporting scheme, health information is aggregated via a bottom up hierarchy converging at the national level to provide evidence-based support for new health policies and eventual policy dissemination of national health strategic plans back down to the districts. The challenge with this hierarchal reporting system is that information often travels slowly between various levels of government, and is often non-responsive to the individual community level needs. Additionally, community level information provided by CHWs is attenuated as a result of aggregation of sub-district level data to district and provincial centers (Freifeld et al. 2010). Furthermore, the current reporting system is labored and time intensive during outbreak investigations (Kamanga et al. 2010). It is often difficult to evaluate the overall CHW contribution towards disease identification, treatment of illnesses, and in reducing disease burdens.

The use of ICT devices for management of CHW registries containing disaggregated community health data and GIS to analyze these datasets may provide insights into local spatial variations in disease distributions and treatments. It would also potentially provide feedback on individual CHW performance along with individual needs associated with illness burden and availability of health resources. Furthermore, it would allow for evaluations of CHW effectiveness in addressing community health needs. An exploration of CHW registries using GIS coupled with the existing ICT technology would help provide evidence for local health managers to be more responsive to CHW needs in the future. It may also provide evidence to support policies that allow for more targeted resource allocation of limited health commodities by identifying areas of high need (Kamanga et al. 2010).

Motivation

Zambia is a malaria endemic country. Malaria is fever related illness that is both preventable and treatable, and wherever environmental controls fail, the need for rapid response and effective treatment is imperative. Malaria annually affects more than 4 million Zambians and accounts for 30% of outpatient visits. It is the primary cause of nearly 8,000 deaths annually (UNICEF Zambia 2014). The most vulnerable populations susceptible to malaria are those located in remote or rural areas, pregnant women, persons who are immuno-compromised, and children under the age of 5. These groups are at an increased risk of death attributable to malaria, with 20% of deaths occurring among pregnant women and 35–50% percent of deaths occurring in children under 5 (UNICEF Zambia 2014). Zambia's National Malaria Strategic Health Plan for 2011–2015 cites the scale up of malaria interventions, including prevention and treatment services, as a main focus area.

Zambian CHW serve in the frontlines of health service delivery, which includes malaria treatment and management at the community level, especially in rural settings. CHW are

equipped with the knowledge and skills to treat malaria at its various stages either by referring severe cases to the nearest health facility or by treating less severe cases with anti-malarial drugs (Chanda et al. 2011). Treatment starts with the early recognition of symptoms and immediate action upon the onset of illness. Having access to CHW for appropriate diagnosis and treatment reduces the likelihood of delayed action and intervention in rural communities.

The introduction of rapid diagnostic test (RDT) in Zambia in the last 5 years has led to more efficient and accurate confirmation of malaria cases. This has culminated into timely and more cost effective malaria case management by CHWs, especially among those living within more remote areas (Chanda et al. 2011). Yet despite the availability of RDTs, low rates of routine usage and adherence to diagnostic guidelines for malaria confirmation have been documented (Chipwaza et al. 2014; Kamanga et al. 2010). These findings may be a result of inadequate training of CHWs or stock outs of RDTs within an area as a result of lapses in the health system infrastructure. Stock outs are sudden shortages in availability of healthcare commodities such as medicines, diagnostic tools, and equipment. This could be attributed to a number of factors including poor logistical planning, weak health system frameworks, manufacturer shortages or limits on production, and expiration of products. Stock outs of health commodities in Africa are a common occurrence and in areas where RDT stock outs frequently occur, CHW capacity is reduced to reliance on symptomatic diagnosis for malaria using the onset of fever to determine probable malaria cases (Mayando et al. 2014; Derua et al. 2011).

Symptomatic diagnosis for malaria has often led to inaccurate diagnosis and over-treatment for malaria. Furthermore this has resulted in unnecessary costs and side effects associated with the use of anti-malaria drugs and possible exaggeration of local malaria disease burden (Derua et al. 2011). Equally, non-febrile malaria cases may be under-diagnosed often leading to progression into more severe non-malaria cases, which would then lead to delayed

treatment and death. Most important, incorrectly attributing fever symptoms to malaria increases the risks of CHW missing diagnosis and treatment of other potentially life threating causes of febrile illness, such as bacterial infection, cholera, typhoid, and pneumonia (Chipwaza et al. 2014; Mayando et al. 2014). It has been documented that non-malarial febrile illnesses result in higher childhood mortality across malaria-endemic countries than malaria (Black et al. 2010). The increased likelihood of population level drug resistance over time, which is associated with over prescription of anti-malarial drug treatments, is also a consideration.

These scenarios are more likely to occur, within remote malaria endemic areas where malaria control interventions are poor, RDT use is inadequate or consistently unavailable, and the existence of overlapping febrile illness outcomes is common. It is for these reasons that routine disease surveillance and programmatic evaluation should be performed at the community level for all febrile illnesses and malaria management. The use of GIS to conduct a geospatial assessment of non-malarial disease/illness burdens in malaria endemic areas at the community level could help stakeholders to better understand better the local variations in overall disease burdens.

Additionally, the use of GIS on health data would lead to adequate management of non-malaria fevers and further exploration into causes beyond the generalized assumption of malaria as the primary source of febrile illness in endemic areas (Acestor et al. 2012). Evaluation of CHW-generated data for community health management has the potential to provide greater insight into local spatial variations in disease burdens and illness case management at a local level. This would allow for further exploration into disease/illness incidence for evidence-base support for future policies encouraging targeted distribution of healthcare commodities and resources. Future research could then explore the feasibility and costs associated with such a targeted approach as well as provide evidence to amend existing policies and provide a

framework for implementation strategies. Moreover, areas where strengthening of interventions are needed or where additional CHW trainings are needed for management of non-malarial febrile illnesses could be identified. Resulting to better implementation of programs and additional understanding of factors associated with increased risk of disease burdens and illnesses at the community level.

This thesis describes the application of GIS and spatial scan statistics to assess illness distribution and CHW illness case burdens at the sub-district level in Chongwe, Zambia. This geospatial analysis of CHW registries explored community level illness distribution to identify clusters of three fever related illnesses, malaria, pneumonia, and diarrhea. The analyses served to illustrate the benefits of sub-district area level data analyses in order to provide a better understanding of the extent to which febrile illness cases exists throughout the district. The use of GIS to perform spatial analyses of malaria and other febrile illnesses at the community level opens up opportunities for in-depth exploration into areas where case management can be improved, and diagnostic commodities would have the greatest impact on malaria surveillance and control efforts within this malaria endemic district.

The following thesis chapters will discuss the application of GIS on CHW health registries to seek answers to the following questions:

1. Are CHW health registries an underused source of information about sub-district level standards of care and health outcomes?
2. Can existing CHW health registries be used to create disaggregated, sub-district level, low-resolution geo-datasets suitable for geospatial health data analyses?
3. Can the use of GIS on low-resolution health geo-datasets provide adequate spatial insights into CHW needs and community health outcomes at a sub-district level?

The hypothesis is that CHW registries are an underutilized source of community health information and can be used to identify areas of local disease/illness distribution that may be masked by current data aggregation methods. A secondary hypothesis is that if GIS were to be integrated in mHealth initiatives at the community level, it would have a great impact on disease and illness surveillance at the community level. Timely monitoring of disease/illness variation could occur which would allow districts to be more efficient in outbreak response and strategic in allocation of resources using this technology.

The ensuing chapters are organized as follows. **Chapter 2** provides background information on the geographic area and population distribution of Zambia. It also elaborates on the research questions and hypotheses that guide this study. **Chapter 3** examines the primary data sources and secondary methods of analysis. **Chapter 4** presents findings of geospatial data visualization and spatial cluster analyses. It also discusses the limitations of the findings. **Chapter 5** provides an in-depth discussion of the findings of spatial cluster analyses. **Chapter 6** provides a discussion on the impact of research findings on policy development. It also focuses on final conclusion and suggestions for future studies.

CHAPTER TWO: BACKGROUND

2.1 Geography

2.1.1 Physical

Zambia is a country located between the latitudes of 10° to 18° South and longitudes of 20° to 30° East in the southern region of the continent of Africa. It is a large landlocked country with a land area of 752,618 square kilometers that is bordered by the countries of Malawi, Zimbabwe, Angola, Tanzania, Botswana, Democratic Republic of Congo, and Mozambique. Zambia's terrain is mostly high plateau broken up by small hills and isolated mountain ridges (CIA World Factbook 2013). Most of the land is flat, with an average elevation ranging between 3500ft to 4500ft above sea level. The lowest elevation is in the Southern region. The capital city is Lusaka and has an elevation of 4,265ft above sea level (CIA World Factbook 2013; Zambia Tourism 2014).

Wetlands account for approximately 5% of Zambian land cover. Lakes Bangweulu, Mweru, and Tanganyika are the three most important natural lakes of the country. The total water area for the country is 9,220 square kilometers. Zambia lies on the watershed between the Democratic Republic of Congo and Zambezi River systems. The upper part of the Zambezi River and its major tributaries, of which the Kafue and Luangwa Rivers are the largest, divide the plateau into several large valleys (Zambia Tourism 2014). Communities located on or near these water areas have higher relative risks for vector-borne illnesses such as malaria, onchocerciasis, trachoma, and for water born bacterial or parasitic illnesses like cryptosporidium, giardiasis, and cholera illness outcomes.

Zambia is administratively divided into ten provincial areas, namely: Central, Copperbelt, Eastern, Luapula, Lusaka, Muchinga, Northern, North Western, Southern, and Western provinces (Figure 2.1). At the time of the 2010 census data collection, there were 74 district

areas, 150 constituencies, and 1,430 wards. The government is comprised of Central and Local Governments with Lusaka city serving as the central level of government. Lusaka province, highlighted in green, is where Chongwe district is located and can be seen in the orange area in Figure 2.1.

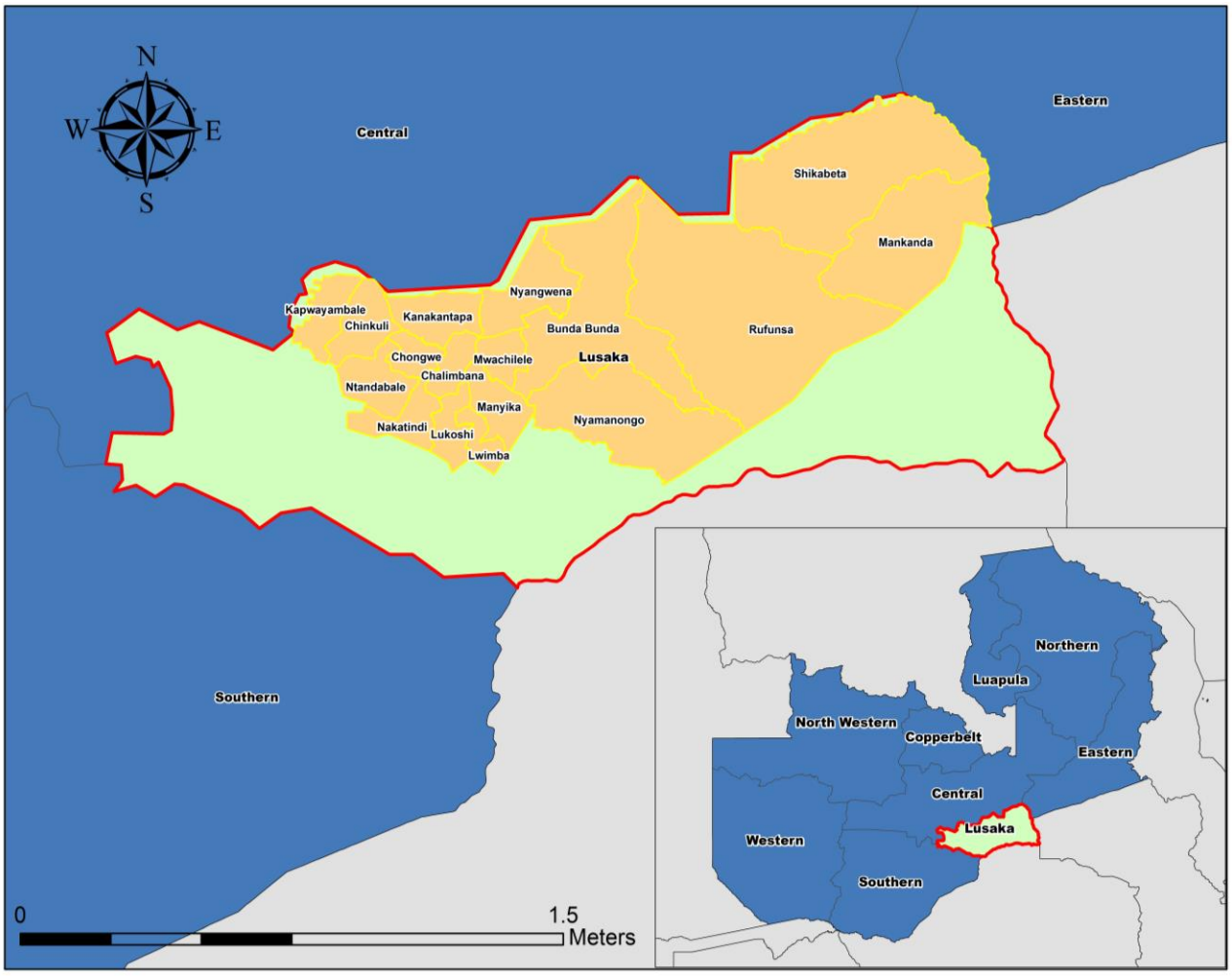


Figure 2.1 Lusaka Province (green) is the location of Chongwe district is located (orange).*

*Muchinga Province is the 10th most newly created Province and is not represented in the inset map.

2.1.2 Climate

Zambia has a tropical climate, with most of country classified as humid subtropical or tropical wet and dry, which is modified by elevation. There are two main seasons, the rainy season and dry season. The rainy season typically begins between November/December and

extends through to March/April, corresponding with the summer season. The dry season typically occurs between the months of May and August, corresponding with the winter season. The dry season is divided into hot and cold dry seasons, with the hot season occurring between the months of September and October/November, and the cold season occurring between May and August. The average monthly temperature throughout the country is 68°F, with higher temperatures occurring in lower elevation areas and valleys (Zambia Tourism 2014). Changes in climate and temperature have been known to be associated with increased incidence of malaria, cholera, respiratory infections, dysentery, and lymphatic filariasis in Zambia (Kasali 2008; Slater and Michael 2012). For many vector-borne diseases, climate is a key factor for disease prevalence. Changes in the raining season have resulted in more incident cases of malaria illnesses in many district areas.

2.2 Population

The total population estimate for Zambia is 13,092,666 and 2,513,768 estimated households. The country is sparsely populated with a population density of 17.4 persons per square kilometer. The most densely populated area is Lusaka province were there are 100 persons per square kilometer. The population is distributed disproportionately between rural and urban areas with approximately 60% living in rural areas. The estimated number of people living in the rural areas is 7,919,216 and in the urban areas it is estimated to be 5,173,450. The population by gender is estimated to be 6,454, 647 for males and 6,638,019 for females (CSO Census Report 2010).

Agriculture is the most common source of income and livelihood, which constitute 85% of the labor force. The mining and industry sectors make up 6% of the labor force and services industry approximately 10%. Sixty-five percent of the population lives below the poverty line, earning less than $1.25USD a day, much of which occurs in the rural areas. The level of poverty

in the rural areas is three times higher than in urban areas, which are estimated at 77.9% and 27.5% respectively. (CSO Census Report 2010).

The number of physicians available per patient is .07 physicians for every 1,000 patients. Physicians are medical doctors, specialists, and general practitioners. The World Health Organization (WHO) estimates that fewer than 2.5 healthcare workers, which include physicians, nurses, and midwives per 1,000, would be insufficient to adequately meet primary healthcare needs. The estimated percentage of health expenditures is 6.5% of Zambian Gross Domestic Product (GDP). In 2010, the Zambian health system had a total of 1,883 health facilities, six tertiary level hospitals, 21 general hospitals, 85 district hospitals, 1,495 urban and rural health centers, and 275 health posts (CIA World Factbook 2013; CSO Census Report 2010; WHO 2012).

Like most of the African continent where 62% of the African population is under the age of 25, Zambia has a young population, with 45% of the population under the age of 15. The median age in Zambia is 16.9 years with the average life expectancy of 52 years. The estimated population under the age of five is 2,252,748 with a mortality rate of 138 deaths per 1,000 live births. The crude birth rate is 35 per 1000 people present at mid-year. The infant mortality rate in Zambia is 76 deaths per 1,000 live births and the child mortality rate is 62 deaths per 1,000 live births. The major causes of child mortality are infectious diseases like pneumonia, diarrhea, malaria, and measles and are common in Zambia. The mean age of a mother's first birth is 19 years old. The population of women of childbearing age (15-49 years) is estimated to be 2,822,635. The national maternal mortality rate is 86.1 deaths per 100,000 live births (CSO Census Report 2010).

2.3 Disease Burdens

Seventy-five percent of all reported deaths in Zambia are attributed to illnesses and disease. Additionally, the major cause of disability in Zambia is also attributed to disease. According to 2012 estimates, 12.7% of adults aged 15 to 49 are living with HIV/AIDS (UNAIDS 2012). Zambia, like most African countries, has a generalized HIV/AIDS epidemic and has made great progress in effective treatment and distribution of antiretroviral therapies. This has resulted in a steady decline in AIDS related deaths in the last decade. This steady decline has led to a shift in focus to address other high burden diseases and illnesses in Zambia. Malaria, Pneumonia, and Diarrhea are among the leading causes of death in sub-Saharan Africa, especially for children under the age of 5.

The overall disease burden in Zambia can be divided into two classes, Non-communicable (NCD) and Communicable diseases (CD). Non-communicable diseases are defined as diseases or conditions that are non-infectious and non-transmissible among people. These may be chronic diseases that have a long duration and slow progression or conditions that are short duration and rapid progression. NCDs include many types of cancers, heart disease, and diabetes, etc. For CDs, these are any condition that is transmitted to a person either directly or indirectly from an infected person or animal. This also includes transmission through an external agency such as intermediate animal, host, vector, or inanimate environment (Annual Health Bulletin 2011). Table 2.1 lists the leading NCD and CD burdens for all age groups in Zambia.

Table 2.1 Leading Causes of Death in Zambia for Non-communicable and C Communicable Diseases

Non-communicable Diseases	Communicable Diseases
Hypertension	Malaria
Asthma	Respiratory Infection (RI): (Non-Pneumonia)
Epilepsy	Diarrhea: non-bloody
Herpes Zoster	Muscular skeletal and connective tissue non-trauma
Cardio Vascular Disease	Trauma other injuries wounds
Diabetes	Digestive System Non-Infectious
Karposi Sarcomas	Respiratory Infection (RI): Pneumonia
Cervical Cancers	Eye Diseases Infections
Breast Cancers	Skin Disease Non-Infectious
	Dental Carries

Source: 2011 Annual Health Statistics Bulletin, HMIS Dataset

2.3.1 Malaria

Epidemiology

Globally it is estimated that 3.4 billion people are at risk of malaria. In 2013, there were 104 countries and territories in which malaria was considered endemic (World Malaria Report 2013). Pregnant women and children under the age of 5 are the population groups with an increased risk of developing severe malaria outcomes and death. Estimates by the WHO state that approximately 207 million cases of malaria and nearly 627,000 deaths occurred globally in 2012. The majority of these cases (80%) and deaths (90%) occurred on the continent of Africa, with most deaths (77%) occurring among children under 5 years of age (World Malaria Report 2013).

Malaria is caused is caused by five species of parasite that affect humans. Of the five, the species *Plasmodium falciparum* and *Plasmodium vivax* are the two most important and *P. falciparum* is responsible for the deadliest form of malaria. *P. falciparum* is the predominant species in Africa, however *P. vivax* has a wider distribution. *P. vivax* has a wider distribution as a result of its ability to develop in the *Anopheles* mosquito at lower temperatures, survive at higher

altitudes, and exist in cooler climates (World Malaria Report 2013). *P. vivax* also has a dormant liver stage, which allows it to survive during winter periods and when the Anopheles mosquitos are not present to carry out transmission. In areas outside of Africa, infections with *P.vivax* are more common than infections due to *P. falciparum.*

Table 2.2 Epidemiology of prevalent malaria species in Zambia

Life Cycle	P. falciparum	P. vivax
Minimum temp needed for maturation in the mosquito	Lowest temp 60.8°F	Lowest temp for cycle to be complete 59°F; Lowest temp for parasite survival, 50°F for two days
Dormant liver stage	No	Yes
Gametocytes	Appear after asexual blood stage is established	Appear at time of asexual blood stage often before clinical symptoms
Disease		
Severity	5% of cases develop into severe illness; responsible for majority of deaths	Risk of severe disease not firmly established
Relapse	No	Yes
Asymptomatic Outcomes	Common	Very Common
Diagnosis		
	Blood film, RDTs and PCR for blood stage	Blood film, RDTs and PCR for blood stage
		No test for dormant Liver stage
Treatment		
Blood stage	Artemisinin combination treatment (ACT) recommended	Chloroquine still efficacious in most areas
Gametocytes	Need single dose Primaquine, Artemesinins have some effect	Sensitive to blood stage treatment
Liver stage		14 days of Primaquine

Source: WHO World Malaria Report, 2013

Malaria Control and Prevention

In endemic countries, where the risk of exposure to malaria is high, great efforts have been made to develop effective malaria control policies and strategies. This has involved the

establishment of well-funded National Malaria Control Programs (NMCP), tailored national and regional strategies, applied and operational research initiatives, and collaboration among multinational partners within the malaria and development communities.

The World Health Organization is a multinational collaborator and primary facilitator of many global malaria control standards and initiatives. The WHO has developed malaria control guidelines and recommends a multi-faceted strategy for malaria control and elimination, which includes prevention therapies, diagnostic testing, Artemisinin-based combination therapies, strong malaria surveillance, and vector control (World Malaria Report 2013).

Vector control has been an important strategy of many malaria control programs. The main objectives of vector control strategies are to lower the intensity of malaria transmission by shortening the lifespans of local mosquito populations; and to reduce, or prevent altogether, the interaction between mosquitos carrying malaria causing parasites and humans. Nearly all vector control strategies involve the use of pesticides and chemicals designed to achieve one or more of the main objectives for vector control strategies. Of the different strategies available for vector control, the most successful are indoor residual spraying (IRS) and insecticide-treated nets (ITNs), including long-lasting insecticide treated nets (LLINs).

Malaria prevalence in Zambia

Malaria is the leading cause of morbidity for all age groups in Zambia affecting more than 4 million people annually. Zambia has made great efforts in improving the extent of malaria control services offered throughout the country. As a result, the government has identified malaria control interventions, both treatment and prevention, as a major public health priority area for the National Health Strategic Plans (National Malaria Strategic Plan 2011–2015).

Between the years of 2009–2011, Eastern Province had the highest incidence of malaria. During this same period, there was a reduction in malaria incidence from 821 per 1,000 persons

in 2010 to 772 per 1,000 persons in 2011 (2011 Annual Health Statistical Bulletin). According to the 2012 Zambia National Malaria Indicator Survey, malaria parasite prevalence was 14.9% with more parasite prevalence occurring in children (22%) living in rural areas. Children in rural areas were more likely to suffer from fevers than those in urban areas (29.3% and 13.2%, respectively). Over 4.7 million cases of malaria were treated at health facilities, with 6.1 million rapid diagnostic test kits distributed to health facilities in 2012.

Malaria Diagnosis and Treatment

The introduction of rapid antigen-detection diagnostic tests (RDTs) into many national malaria control programs, has made malaria case detection more reliable, and has improved the timeliness of appropriate treatment interventions. Prior to the introduction of RDTs malaria diagnosis was clinical (symptom-based) and was proven to be quite unspecific as to what symptoms were reliable for diagnosis (Msellem at al. 2009). The presence of fever was the most commonly used symptom for malaria clinical diagnosis and treatment was given for malaria. Additionally, children with fever who were not confirmed to have malaria were also given antimalarial medications, as precautionary measures for malaria illness. This practice referred to as presumptive treatment has become less appropriate given the increased availability of low cost RDTs and lower transmission rates of malaria.

Fever is often an overlapping clinical symptom for malaria, pneumonia, and diarrheal diseases, which are all common illnesses in resource poor settings or developing regions. In malaria endemic areas, fevers in children under the age of five are often attributed to malaria illness (Kazembe et al. 2007, Amexo et al. 2004, d'Acremont et al. 2010). Children under the age of five account for ninety-percent of all malaria deaths. In malaria endemic areas, the onset of fever in children is often attributed to malaria, regardless of clinical confirmation, and treatment is therefore given for malaria (Rao, Shellenberg, and Ghani 2013). This method of

symptomatic diagnosis has led to misdiagnosis and overtreatment of malaria in many endemic areas (Amexo et al. 2004). Misdiagnosis and over treatment of drug therapies has been known to reduce the effectiveness of drug treatment and is a major cause of drug resistance.

For children who are misdiagnosed as malaria, the true cause of fever remains untreated resulting in adverse outcomes, potential drug resistance, prolonged and more severe illness, or death (Amexo et al. 2004, Hume et al. 2008). The development of RDTs was an important innovation for malaria case detection. It created opportunities in malaria endemic areas and resource poor or remote settings to have access to parasitological confirmation of potential malaria cases (Hamer et al. 2007).

The WHO has taken steps to acknowledge the mismanagement and under-diagnosis of non-malarial febrile illnesses, with high mortality rates, among children. The World Health Organization and the United Nations Children's Fund (UNICEF) released a joint statement in 2006 encouraging the practice of integrated case management for febrile illnesses among children less than 5 years of age. The encouragement by the World Health Organization and the United Nations Children's Fund reflected a shift from previous standards in patient management that was focused solely on febrile malaria cases. The new standard for patient management was a shift towards more concurrent approaches that focused on other common non-malaria febrile illnesses like pneumonia and severe diarrhea among children less than 5 years of age (Chanda et al. 2009; Young et al. 2012; World Malaria Report 2013).

According to a review conducted by the Child Health and Epidemiology Reference group, it is estimated that a 70% reduction in mortality from pneumonia in children could result due to community efforts to manage all cases of childhood pneumonia (Young et al. 2012). The review also estimated the use of oral rehydration salts and zinc, at the community level in treatment for diarrheal illnesses, could prevent 70-90% of deaths caused by acute watery diarrhea

(Young et al. 2012). These findings illustrate the impact of addressing non-malarial febrile illnesses. The use of RDT kits has led to more accurate diagnosis of malaria, however for cases of febrile illness confirmed not to be malaria, the cause of fever is difficult to establish (Chanda et al. 2009; White et al. 2012).

The use of RDT also improved drug treatment management and reduced the early onset for drug resistance. Early drug treatment involved use of monotherapies either with Chloroquine or Sulfadoxine-pyrimethamine. As a result of common usage, misdiagnosis and overtreatment, drug resistance quickly emerged for malaria treatment using Chloroquine and Sulfadoxine-pyrimethamine. This emerging resistance prompted many African countries to change treatment policies to Artemisinin-base combination therapy (ACT) as first line treatment for non-severe malaria cases (Msellem et al. 2009). The combination of RDT and ACT was an important strategic shift towards reducing morbidity and mortality associated with malaria in endemic regions. In Zambia, the malaria control program focuses on the use of Coartum as first-line drug therapy. Sulfadoxine-pyrimethamine is the primary drug treatment for non-severe malaria in children, and Quinine is prescribed to the most severe malaria cases.

ACTs have become a mainstay in malarial drug treatment therapies and involve the concurrent use of two drugs, one of which is Artemisinin based, to slow down the evolution of parasite resistance. Reports of ACT resistance have been published and are cited in Hastings (2011). These findings have caused some concern in the malaria research community and spurred research into the causes of resistance. Hastings (2011) comments, that combining various drugs with Artemisinin has delayed progression into more severe drug resistance that would have occurred with monotherapies. Hastings (2011) suggests that emphasis now should be placed on good data capture accompanied by efforts to improve diagnosis and drug compliance to maintain the efficacy malaria treatments.

2.3.2 Pneumonia Infections

Pneumonia is the single largest cause of death in children under 5 years of age and kills an estimated 1.1 million children annually, which is more than malaria and tuberculosis deaths combined. Pneumonia is an acute respiratory infection that affects the lungs and can be caused by viruses, bacteria, or fungi. It is a preventable disease through immunization, adequate nutrition and control of and protection from environmental factors (WHO "Pneumonia 331" 2013).

Pneumonia can be spread through the air and through blood contact. For viral and bacterial pneumonia, clinical symptoms are similar, with viral pneumonia having more explicit symptoms than bacterial pneumonia. In children under the age of 5, who have a cough and/or difficulty breathing, with or without fever, pneumonia is diagnosed by the presence of either fast breathing or constricted chest movements during breathing inhalation. Children with weakened immune systems or who are malnourished are at increased risk for developing pneumonia infection (WHO "Pneumonia 331" 2013).

Most cases of pneumonia can be treated with antibiotics, with most requiring oral antibiotics, which are often prescribed at a health center. In Zambia, acute respiratory infection is among the top five causes of morbidity and hospital admissions for both adults and children (Annual Health Bulletin 2011). Between the years of 2009 to 2011, the pneumonia incidence rate steadily increased from 29.3 to 35.9, respectively for number of new cases per 1,000 persons. Eastern province had the highest incidence rate during that same time period and Lusaka province had the lowest incidence rate (Annual Health Bulletin 2011).

2.3.3 Diarrheal Diseases

Similar to pneumonia, diarrheal diseases are both preventable and treatable. Diarrheal diseases are the second leading cause of death in children under five years of age, accounting for nearly 760,000 deaths annually. Diarrhea is a leading cause of malnutrition among children (WHO "Diarrhea 330" 2013). Annually, there are nearly 1.7 billion cases of diarrheal diseases globally and diarrhea is a leading cause of malnutrition in children under five years. Many of these cases can be attributed to poor sanitation and hygiene in addition to a lack in access to safe drinking water (WHO "Diarrhea 330" 2013).

The clinical definition of diarrhea is the passage of three or more loose or liquid stools per day or more than normal frequency of passage for an individual. There are three clinical types of diarrhea, acute watery diarrhea, acute bloody diarrhea, and persistent diarrhea. Diarrhea is usually a symptom of infection caused by virus, bacteria, or parasitic organisms in the intestinal tract. Infection is spread from person to person due to poor hygiene and may also be spread through contaminated food and drinking water (WHO "Diarrhea 330" 2013). Diarrhea due to infection is common in many developing regions. The most severe symptom of diarrhea is extreme dehydration where water and essential electrolytes (sodium, chloride, bicarbonate and potassium) are lost and not replaced as a result of watery stools, vomit, sweat, urine, and fast breathing. Death can follow severe dehydration if fluids and electrolytes are not replaced.

The prevention of diarrheal outcomes involves interventions that include accesses to safe drinking water, improved sanitation and hygiene systems and practices, and improved hand washing practices with soap. Treatment for diarrhea can be performed through consumption of an oral rehydration solution (ORS) packet containing clean water, sugar and salt, and with zinc tablets. Globally, 780 million individuals lack access to clean drinking water and 2.5-billion lack

access to improved sanitation conditions (WHO “Diarrhea 330” 2013).

In Zambia, diarrhea is among the top ten major causes of morbidity and dehydration. During the time period of 2009-2011 diarrhea incidence were 72 cases and 86 cases per 1,000 persons respectively, reflective of a steady increase for diarrheal outcomes over time (Annual Health Bulletin 2011). All provincial areas had a steady increases in diarrheal outcomes during this time with Southern and Northwestern provinces recorded the highest incidence rates. Lusaka and Northern provinces had the lowest incidence in 2011 with 74 cases and 68 cases per 1,000 persons, respectively.

2.4 Mobile Health (mHealth) Applications

Mobile health innovations have helped to improve the overall quality of care in countries like Zambia and other Sub-Saharan African countries. Mobile health (mHealth) is a subset of information communication technology (ICT) applications, which relies on the use of mobile phones, personal digital assistants (PDAs), tablets, smart phones, and other mobile devices to manage and share health information and data.

Kamanga et al. (2010) used mobile phones and GIS to detect potential outbreaks of malaria to identify locations where parasite reservoirs are likely to occur for targeted interventions. The study occurred in Southern Province of Zambia in Choma and Namwala districts. This study area had reliable mobile network reception and nearly all residents and at 12 out of 14 rural health centers (RHC) had access to the network.

The CHW entered routine health data into the health center registry daily. Once a week, data for positive malaria diagnosis and total number of RDT used per week by SMS to researchers at the Malaria Institute in Macha from each health center. Each SMS transmission contained information that included the RHC name, name of transmitting nurse, number of RDT used during the week and number of positive diagnosis with RDT. The data were then entered

into a spreadsheet and sent to District Health hospitals, National Malaria Control Center (NMCC) and John Hopkins University in Baltimore, Maryland. Each weekly spreadsheet consisted of data from 12 of the 14 RHC, two RHC were omitted from the study due to poor phone access. The data was stratified by RHC into two separate ecological zones, flood plain area and an area of increasing aridity and elevation. Comparing SMS texts with actual registry data at each RHC validated data and any discrepancies were investigated and ameliorated. The participants were compensated for the use of their mobile phones and provided a modest stipend.

GPS coordinates were also collected for each RHC and plotted on a Landsat map that included elevation data and other features. The weekly malaria incidence rate was calculated for each RHC catchment area. These methods illustrated the plausibility of mobile phone usage for routine surveillance activities and that the timeliness and efficiency in resource allocation for malaria control operations could be improved in doing so. The use of GIS also served to identify areas where patterns of RDT usage affected the overall cost effectiveness of diagnostic testing for malaria.

The authors concluded that through this case study, there is evidence that use of mHealth approaches for reporting critical data would provide opportunities rapid deployment of specific interventions and drugs to prevent malaria transmission. The network of CHWs recruited into the program had been receptive and highly cooperative, and regular feedback of data and information was conducted. The use of mobile phone system was effective and inexpensive and broader implementation would allow for more timely communication between RHC and central health authorities. All expenses and reimbursements were handled through the SMS system.

In Tanzania, a similar study evaluating SMS and mobile phone usage was conducted (DeRenzi et al. 2012). In this study, researchers evaluated the impact of SMS reminders to improve the timeliness of routine CHW visits to homes of clients located in their communities.

In Tanzania, CHW success has been tied to routine home visits, which can be difficult to maintain as a result of shortages in available equipment, supplies, and reliable transportation.

DeRenzi et al. (2012) introduced mobile phones to CHW as a tool for community health data management. They initially began with a pilot study in Dodoma, Tanzania for 9 weeks, to evaluate the impact of SMS reminders to improve CHW performance. This was then followed by two larger studies implemented in Dar es Salaam, Tanzania over a 9-month period with 87 CHWs. The authors introduced a reminder system that augmented existing supervision structures within CHW programs. This system sent proactive reminders to CHWs the day before and the day of scheduled routine visits.

Based on results and feedback from the pilot study, the SMS reminder system was modified in two ways. First, in the pilot study, the first SMS reminder was sent after working hours, in the evening of the day that the CHW's visit was due, which was too late for an on time visit. This was addressed in the Dar es Salaam studies with the introduction of *proactive* reminders, with the earliest SMS reminder sent the day *before* a visit is due. Secondly during the pilot study, if a referral was not closed during the pilot, the system continued to send SMS messages but had no capabilities of determining *why* the CHW was not reporting a follow-up visit. This led to the addition of the SMS reminder escalation to the supervisor stage for intervention, implemented in the two larger studies in Dar es Salaam.

The strengths of this study can be found in the use of a control and intervention group, a pilot phase that informed a larger study, and comparison of two methods illustrating the impact of structured supervision. The limitations of the study can be found in the accuracy of the findings with actual follow-ups and ground trothing. Unlike, Kamanga et al. (2010), which verified mHealth data with data available at each RHC, DeRenzi et al. (2012) did not verify if CHWs actually visited clients. Logistically this was not feasible, however upon Supervisor

escalation, CHW would call or SMS an automated number and provide reasons why to supervisors, which provides some evidence that visits actually occurred. The authors note that future applications should make use of GPS and lower cost mobile phone technology to address this limitation.

Braun et al. (2013) provide a systematic meta-analysis review of peer-reviewed literature focused on evaluating the impact of mobile technology on community health workers' utilization in healthcare management. They reviewed 25 research articles published between January 1, 2000 and June 30, 2012 to identify opportunities and challenges for strengthening health systems in resource-constrained settings. They analyzed and coded each article for various topics then compiled findings into tables according to key topics. The 25 articles reviewed involved 28 unique studies and most reported on projects in developing countries. There were more programs operating in rural areas than in urban areas, with a few operating in both. A broad range of health issues was addressed with the most common topic areas being sexual, reproductive, and maternal and child health or which more than half focused specifically on HIV/AIDS. Other key topic areas were tuberculosis and malaria. In the articles reviewed, mHealth technology was most commonly used for data collection, decision support, alerts and reminders, information on demand, and as a facilitation tool for CHW activities associated with field based research and direct provision of medical care.

The authors concluded that more analyses evaluating the cost effectiveness of mHealth utilization by CHW would be helpful for program staff and policy makers. Future research should also focus on utilization of qualitative data collection methods in order to better understand how mHealth tools may be improved and adapted to local CHW contexts and performance. They also note limitations of their review, citing the number of articles reviewed unlikely reflects the scope of mHealth projects being implemented. They go on to state that most

projects were focused on deployment of mobile tools to enhance service delivery rather than scientific interests, it is likely that many mHealth projects are not reported in published literature.

In another meta-analysis review conducted by Nhavoto and Gronlund (2014), the authors, unlike Braun et al. (2013), sought to identify applications of mHealth initiatives that integrate mobile technologies and GIS together in concurrent application. Through their analyses they identified how mobile technologies and GIS applications have been used independently as well as in combination in healthcare information infrastructure as a means to provide a basis for data analysis and decision-making support.

The authors selected a total of 271 articles centered on use of mobile technologies and GIS in improving healthcare, 220 focused on mobile technologies and 51 GIS. Most of the articles reviewed involved studies implemented in developed countries, most notably the United States. The articles were then sorted into six predominant themes: treatment and disease management, data collection and disease surveillance, health support systems, health promotion and disease prevention, communication to or between health care providers, and medical education. Applications of GIS technology were sorted into four predominant themes: disease surveillance, health support systems, health promotion and disease prevention, and communication to or between health care providers.

Although most of the articles reported positive results, Nhavoto and Gronlund (2014) identified research gaps. First, the overwhelming majority of publications reported positive results, suggesting that unsuccessful applications were under reported. Secondly, evaluated publications focused on effects and often failed to discuss implementation efforts, creating barriers for scale-up into broader applications and adaptation for various settings. The authors cite this as an important research gap and an indication of a more serious problem. Despite all of the successive outcomes, many of the interventions were performed as pilot projects or small-

scale projects, with very few scaling up to larger scales. A possible explanation suggested by the authors was the challenge of aligning many more actors and stakeholders, standardizing data, upfront investment in digitization of data, and legal, economical, and practical constraints regarding communication in larger scale implementation. Research into implementation feasibility may need to be conducted in these areas before large-scale interventions occur. Finally, the findings of this review also concluded that little integration between GIS and mobile technologies occurs in many of the mHealth interventions reviewed. Nhavoto and Gronlund (2014) conclude that in order for mHealth information processes to be most effective they must integrate different kinds of existing technologies. The increasing development of mobile apps provides additional opportunities for integration of mHealth health application and management tools.

Although Nhavoto and Gronlund's (2014) criticisms are valid, there are often numerous barriers to GIS integration. The use of GIS in itself can be quite costly and integration with other technologies may be even more expensive. In addition to this, the training costs both in time and funding, are added barriers to GIS utilization in many developing regions and health systems. Mobile phones with GPS or GIS components are also costly despite the increasing availability of smartphones in developing countries. Many smartphone applications require reliable Internet network access and this can prove difficult in remote areas. Numerous open sourced GIS software and platforms have emerged, such as QGIS and R, reducing and even eliminating some of these barriers.

In the public health sector, SaTScan v9.3 is an open-sourced software developed by Martin Kulldorff, that analyzes datasets for spatial, temporal, and space-time clustering and has been widely used to perform hot spot analysis on health data. It was designed for geographical surveillance of diseases through spatial or space-time disease cluster detection and assesses

statistical significance of identified clusters. SaTScan v9.3 can be used on discrete or continuous data for spatial scan statistics through a number of scanning options for various data types. Epidemiologists in developed and developing regions have used SaTScan v9.3 to detect disease clusters for cancers, malaria, and TB among a number of other diseases (Coleman et al. 2009; Zhao et al. 2013, Sherman et al. 2014). The low cost and ease of use, allow for SaTScan v9.3 to be an alternative option to complicated propriety software for GIS analyses, and provides an opportunity for resource strapped health systems to conduct geospatial analyses on their health data.

Nhavoto and Gronlund's (2014) claimed that mHealth information processes would be most effective if they integrate different kinds of existing technologies. For mHealth programs in Zambia, given the availability of low cost mHealth and GPS devices, what would be the impact of integrating existing mHealth devices with GPS capabilities and programming? Would the integration of GPS and mHealth devices be more advantageous for CHW community disease/illness case management and treatment? Given the positive outcomes of previous mHealth programs and initiatives in Zambia on CHW case management and performance, it is predicted that the integration of mHealth devices with GPS would prove to be more advantageous. It would allow for data coupling with GIS for in-depth spatial data analyses at the sub-district level. Community areas could be explored for non-random disease variability and clustering. Identified hot spots or clusters would benefit from more targeted interventions. This thesis will aim to substantiate this hypothesis by exploring the potential benefits of integrating existing mobile technologies used by CHW with GPS capabilities for routine health data management and activities. The proceeding methods and analyses describe the use of GIS on CHW community health registries to assess illness distribution at the sub-district level in Chongwe, Zambia.

To explore this potential impact, a geodataset will be created from paper-based CHW registries to model the expected output of a GPS enabled mHealth device. Spatial analyses using SaTScan v9.3 on geo-referenced CHW health registry will seek to identify clusters of febrile illness related malaria, diarrhea, and pneumonia at the community level. The assumption is that these registries can highlight areas where diagnostic commodities and scarce health care resources would have the greatest impact on disease/illness surveillance, control, and treatment efforts within the district.

This thesis will build upon a previous mHealth program for the proceeding secondary data analyses and will explore health registries managed by CHWs in Chongwe during the Grand Challenges Canada (GCC) mHealth feasibility study. SaTScan v9.3 will be used for spatial statistics on a sub-district administrative area to evaluate and identify non-random illness distribution in Chongwe.

2.5 GCC mHealth Study in Chongwe District, Zambia

2.5.1 Overview

The GCC mHealth study was a Zambian IRB approved community-based project conducted between the months of January and October 2013 (IRB# 00006464; IORG#00005376). It was a pilot study that evaluated the feasibility of mobile phone usage among CHWs for community health data management in the rural district area of Chongwe, Zambia.

The program was a feasibility study to assess the effect of ICT devices on community health performance in Zambia (Chanda et al. 2009; Chanda-Kapata et al. 2011). Furthermore, the study aimed to explore the effects of mobile health and ICT devices on CHW retention and on the quality of health services provided at the community level. The researchers primary objectives sought to answer the following questions:

1. What would be the impact of ICT devices on community health services and CHW motivation and retention in rural Zambia?

2. What would be the impact of ICT use for community health services on the health status of communities in Zambia?

3. What is role would ICT devices have on the capabilities for health center staff to monitor CHWs?

A mixed methods approach of both qualitative and quantitative data collection methods was used. Seventeen health facilities participated in the study and were randomly assigned into a control or intervention group to determine the effects of ICTs on CHW retention rates. Eight health facilities were randomly selected and placed in the control group and 9 in the intervention group. The health facility staff in each facility included primary nurses, environmental health technicians, and clinical officers who acted as supervisors to CHWs.

A preliminary survey was conducted to determine the perceptions of ICT device utilization among community and professional health workers. Successful utilization by CHWs, promised to improve information management, incentivize CHW retention, and improve community education and services.

2.5.2 Study Area

The location of the study area took place in an area called Chongwe, a rural district are located approximately 35 Km from the capital city of Lusaka in Lusaka Province. It is an area with high malaria prevalence within the communities. Agriculture and farming are the main land uses and source of income by residents. Chongwe's climate is arid and it is nestled in valley region surrounded by numerous hills. This district area was selected as the study location because of previous work with CHWs on malaria case management (Chanda et al. 2011). It is also a sentinel site for malaria surveillance for the region.

The district is divided into rural health facility catchment areas, which are further sub-divided into CHW zones. The health facilities that fell within the catchment area zones of each participating CHW were also included in the study. Each CHW reported to one health facility within their catchment area, so multiple CHWs may report to one health facility. Health facilities that did not have CHWs were excluded from the study.

2.5.3 Population

The population estimates for Chongwe district were 192, 303 (2010) and for each CHW zone the population ranged from 200–1000 individuals. Each CHW zone had a number of farming villages located within their respective catchment areas. The district had 40 health facilities with a total of 377 health workers, of which 245 were health professionals, 116 were daily employees, and 16 administrators. The administrative staff was based in the District Health Office. All residents who lived within the catchment areas of participating CHWs were included in the study. Resident participants were any individuals who sought health services from the CHWs, both in the intervention and control areas, during the time period of January to October 2013. The primary health services provided were mainly for malaria, pneumonia, diarrhea, and health education.

2.5.4 Community Health Volunteers

The district had a total of 88 CHWs and CHWs who resided within their respective zone and were currently not involved in any other research program were included into the GCC study. The CHWs who participated had been previously involved in earlier health research studies in the area and therefore had an existing relationship with researchers (Chanda et al. 2011). The professional health staff for each health facility served as supervisors for each CHW.

In the study each community health worker training status was evaluated through a questionnaire along with their currently existing responsibilities. After the baseline assessment,

training began for those selected in the intervention arm on how to use a mobile phone for routine health data collection for program activities. The training involved an overview of the project, ICT devices and their application to health, reporting format and frequency, and how to generally care for and security of the devices (mobile phones and solar chargers). The training for the CHW in the control group was a refresher course for Integrated Community Case Management (ICCM).

The CHW were provided with conditional airtime to add to their mobile phones for the purposes of transmitting weekly SMS summary reports to the central district reporting system. The summaries were stored electronically on the District Health Information System 2 (DHIS2) platform. Additionally, CHWs in both the intervention and control groups kept handwritten registries for patient specific epidemiological information management. These registries served as mobile data backups for each CHW involved in the intervention arm of the program.

2.5.5 Data Collection

The patient level variables collected included patient ID, sex, age, location name, medical complaint, and treatment outcome. The participant information was completed for each visit. On a weekly basis, CHWs composed aggregated summary reports, which were then sent by SMS to the DHIS2 server at the nearest health facility. Health facility staff and central level program staff reviewed this information using password-protected access on the DHIS2 server routinely. The server had restricted access preventing health facilities from viewing health data provided by other health facilities. Monthly monitoring was conducted by review and comparison of copies of the CHW registries to DHIS2 databases.

For the control sites the patient information was collected and entered into the CHW registries only and manually transferred to the nearest RHC for data entry into the DHIS2 server monthly. Health facilities in the control group operated per the usual guidelines.

2.5.6 Outcomes

The study indicated that it was feasible to implement ICT interventions using mHealth devices in rural areas. All of the CHWs and health facility staff in the intervention group were able to use the reporting system effectively. Epidemiological data was available for all levels of care in a timely manner at the intervention facilities. The ICT devices had positive effects on the community health services and retention of CHWs. Perceptions among community members, CHWs, and health facility staff were positive, creating an opportunity for broader implementation of mHealth application for community health initiatives.

Overall, mobile phones as a tool for CHW case management was beneficial. However, there were some issues in regards to the model of the phone and the capacity of the district to respond to the reports received from the various sub-district reporting areas. Additionally this initial study did not integrate the use of mobile phones with other technologies like GIS. This is in line with similar studies noted in Nhavoto and Gronlund (2014). However, unlike most mHealth programs, spatial thinking was involved in the development of the program. One of the secondary objectives was to "develop real-time epidemiological maps of notifiable diseases." In order to achieve this secondary objective, the geospatial data analyst assigned to this study performed the following methods.

CHAPTER THREE: METHODOLOGY

3.1 Datasets

3.1.1 GCC Dataset

In Zambia, spatially explicit data is not routinely collected for research purposes. Geospatial thinking does not occur during program or study development, resulting in retrospective GPS coordinate data collection activities or the omission of any spatial evaluation altogether due. An added challenge was the lack advanced road network and address systems. Street names, addresses, and zip codes, typically used for geocoding algorithms and programs in more developed regions like the United States, do not extensively exist in Zambia, most especially in rural areas. Without advanced road network and address systems, traditional geocoding methods are often more difficult to perform, and are at present great barriers towards geospatial applications in developing countries. Some well-funded non-governmental organizations (NGO) have overcome this barrier by using GPS devices to obtain coordinate locations. This approach to geolocation data collection can often be expensive and time-consuming for everyone to do. For researchers, the optimal method for geospatial data collection would be to integrate GPS data collection methods, into routine data acquisition activities, during ongoing research programs.

To spatially examine the CHW registries with GIS, the existing GCC database needed to be geocoded, but unfortunately investigators had not collected any GPS latitude or longitude coordinate points for patient or facility locations. Due to limitations in research funding and human resources, it was not feasible to use a GPS unit to retrospectively collect coordinate data points for patients' household locations in Chongwe. As the only member of the data analysis team with GIS knowledge assigned to this study, it was proposed that use of the village names recorded in the registries, may be useful for manual geocoding methods for village coordinate

locations without the use of a GPS device. The feasibility of this alternative approach to coordinate data collection was unknown and had not been documented in any currently available literature.

A de-identified CHW dataset was obtained for this project. The registries were entered into an Excel spreadsheet, the variable names for latitude and longitude were added to the sheet next to the village and health facility variable in the spreadsheet. The name of the village was used as a patient's location and residential address. There were two methods that were attempted to determine the feasibility of manually geocoding villages by name. The first involved the use of Google maps and Google Earth software to enter the village names directly into the search bar to determine the geo-coordinates for each village location. This was unsuccessful due to the duplicate naming of certain villages and cities in Zambia. Some village areas had the same name as other geographic locations in other provinces, resulting in the wrong coordinate locations showing up in searches. It was found that this would not be a feasible method for coordinate identification as a result duplicate location nomenclature.

The second method required collaboration with other researchers and organizations that had done work in Chongwe or near these village areas. Chongwe is a malaria endemic area and is a sentinel site for a number of NGO and governmental malaria research programs and surveillance initiatives. PATH, an American based NGO focused on malaria research and health systems development programs, had conducted numerous studies in the area in previous years. Many of PATH's research activities involved the use of mobile GPS devices and GPS enabled PDAs to collect coordinate locations of local landmark areas and villages for their malaria disease mapping and modeling initiatives. In 2010, researchers from PATH had conducted a study that required the collection of various geolocations within Chongwe district at the community level. The data was later stored in an Access database and included the village names

and corresponding latitudes and longitudes for numerous village locations in Chongwe (Figures 3.1 & 3.2).

A request was made to officials at PATH to have access to this dataset. PATH officials were gracious to share their village location geodataset and it was used as a reference log to identify the village names in Chongwe for the GCC dataset. For each data entry in the PATH log, a GPS coordinate was recorded and a note was included as to where the coordinate points were collected and an indication of if the location was a village or landmark.

After sorting for all of the records in the GCC dataset that had a village name recorded, each record with a village name in the GCC dataset was copied. Using the search function of the Access dataset, village names were manually searched for in search for a matching location in the PATH log. If there was a match, then the geographic coordinates recorded in the PATH log were copied and pasted in the columns for Latitude and Longitude respectively, for the corresponding village in the GCC dataset. This was done for each village location recorded in the GCC dataset. Additionally, a second dataset was obtained from the Central Statistical Office (CSO) that also contained a number of village coordinate locations in Chongwe. The locations were obtained during 2010 census activities, however the dataset was not entirely complete with many of the GPS coordinates missing the naming reference for each coordinate location.

In total of the 7,673 patient records, 2,529 did not have a village name recorded for the patient, and of the 5,144 that had a village name recorded, 3,130 were geocoded using both the CSO and PATH dataset log. A total of 141matching village locations that were identified with corresponding geo-coordinates in the reference datasets.

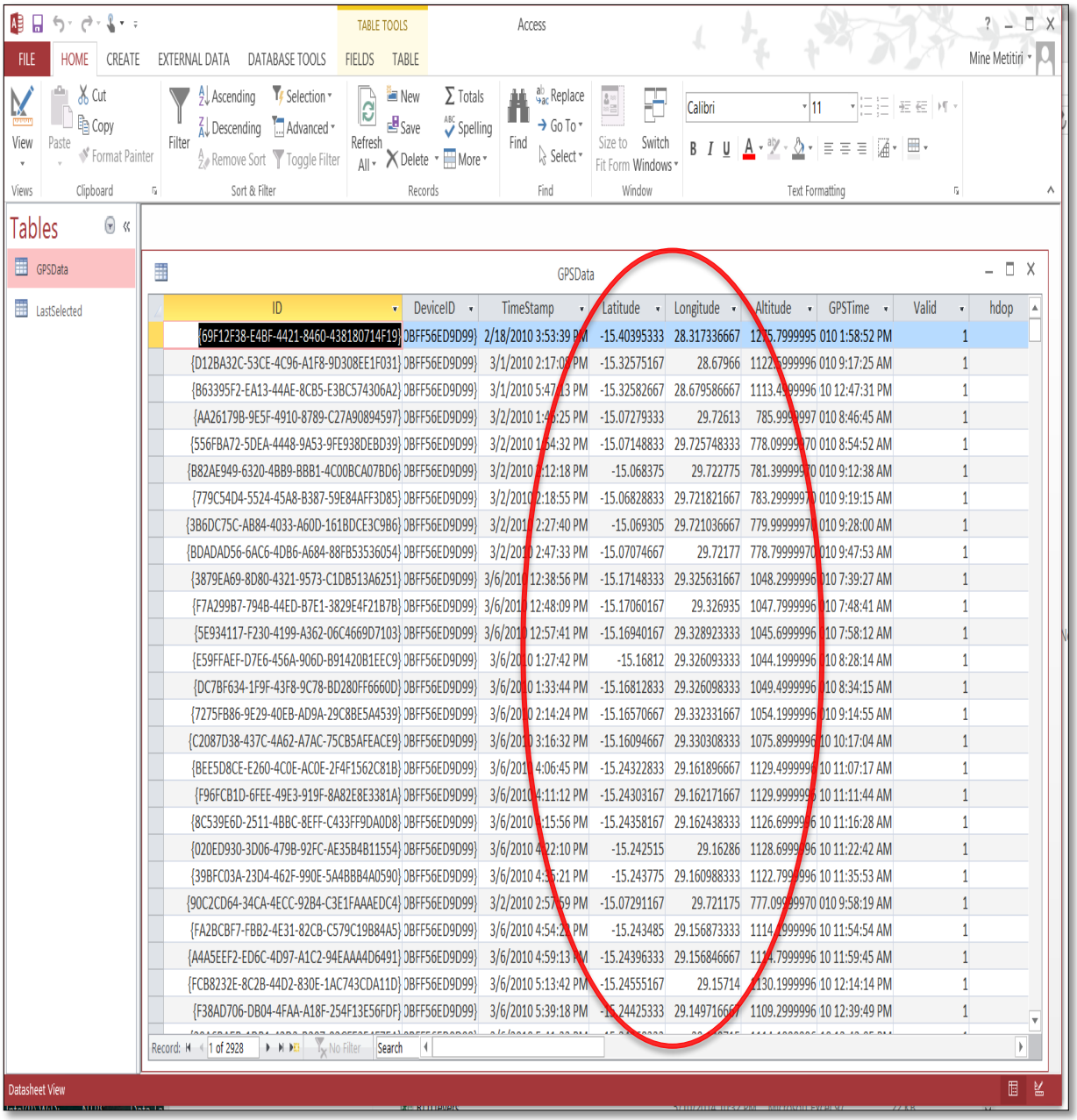


Figure 3.1 PATH log in an Access database containing latitude and longitude coordinate locations for Chongwe villages

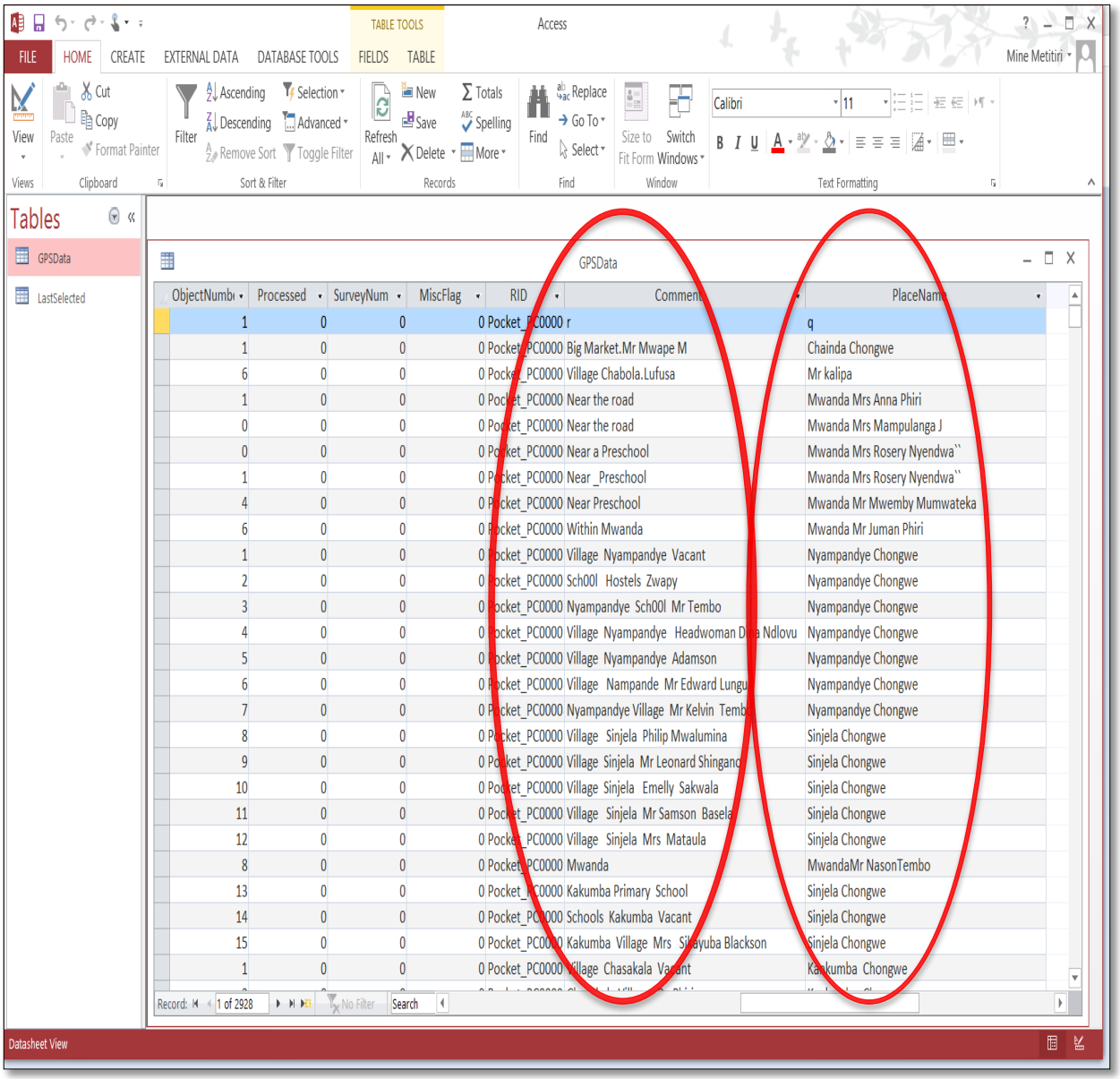


Figure 3.2 PATH log in Access database containing the geocoded village names used matching village names

In addition to geocoding village names and locations, health facility locations (n=17) were also geocoded. This task was made easier by the availability of datasets with health center names and coordinates managed by the CSO and the Zambian Ministry of Health (MOH). After a formal request was made to CSO, the health facility dataset was obtained, along with administrative area shapefiles. Each health center was identified along with the corresponding

coordinates, which were then entered into the corresponding coordinate variable name for each health facility. Of the 17 RHC that participated in the GCC study, only 15 health facilities were geocoded, due to no village location records for the two missing health facilities (Table 3.1).

Table 3.1 GCC Health facilities included in GCC geospatial dataset

GCC mHealth Feasibility Study Dataset: Health Facility Included	GCC Geospatial Evaluation Dataset: Health Facility Included (Yes/No)
Katoba RHC	**Yes**
Mpanshya Hospital/HAHC	**Yes**
Rufunsa RHC	**Yes**
Chinyunyu RHC	**Yes**
Lwiimba RHC	**Yes**
Kasisi RHC	**Yes**
Nyangwena RHC	**No***
Chainda RHC	**No***
Chalimbana RHC	**Yes**
Kanakantapa RHC	**Yes**
Waterfalls RHC	**Yes**
Kankumba RHC	**Yes**
Lukwipa RHC	**Yes**
Kampekete RHC	**Yes**
Mpango RHC	**Yes**
Mwalumina RHC	**Yes**
Shikabeta RHC	**Yes**

*Location not included in geo-dataset due to missing village names for patients in CHW registries

After the coordinate locations were identified and entered into the Excel spreadsheet for both the health facilities and village names, verification on the accuracy of these coordinates was performed using the administrative area shapefiles obtained for Chongwe district from CSO, in a GIS (Figure 3.3). The health facility and village points were input into ArcGIS 10.2 as 'X,Y' coordinate points and were then projected onto the district area shapefiles. The points were validated through visual inspection and all were accurately placed within the Chongwe district area boundary. For every village coordinate location identified (n=141), each location represented multiple patient cases within the village.

P	Q	R	S	AQ	AR	AS
Healthcenter	**Latitude_RHC**	**Longitude_RHC**	**RHC_CA_POP**	**Patient_Address**	**PA_Latitude**	**PA_Longitude**
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Kapuka	-15.345550	28.719527
Chalimbana RHC	-15.386350	28.702283	5025	Mukankaulwa	-15.383392	28.719977
Chalimbana RHC	-15.386350	28.702283	5025	Mukankaulwa	-15.383392	28.719977
Chalimbana RHC	-15.386350	28.702283	5025	Mukankaulwa	-15.383392	28.719977
Chalimbana RHC	-15.386350	28.702283	5025	Mukankaulwa	-15.383392	28.719977

Figure 3.3 GCC spatial dataset Excel spreadsheet with completed health facility and village coordinates

3.1.2 CSO Administrative Area Shapefiles

The next steps required that the points be assigned to lower level administrative areas. The data collected by CHWs for each health facility and village location were collected at the lowest area administrative unit called Standard Enumeration Areas (SEA); however, to determine which SEA code corresponded to each village and health facility location, it was necessary to work first from higher level administrative areas down to the SEA levels. Shapefiles

were requested from CSO for four hierarchal administrative area boundaries for Chongwe, the district, constituency, ward, and standard enumeration boundaries. These were then uploaded into ArcGIS. In looking at the attribute tables, the FIDs were identified for each area and incorporated into the GCC spatial dataset for future joins and relations of tables. The SEAs were available only as FID codes (ex. 501073061031) requiring that to accurately identify each village area within the right administrative boundaries and to the accurate codes, the GeoIDs for the higher levels would have to be obtained prior to SEA assignment (Figure 3.4). Once the GeoIDs were identified for each administrative area level and entered into the GCC Excel dataset for each village location; the file was then uploaded into ArcGIS. Using the GCC dataset, each shapefile was spatially joined to the attribute tables for the district and sub district administrative areas based on the matching FID and GeoIDs.

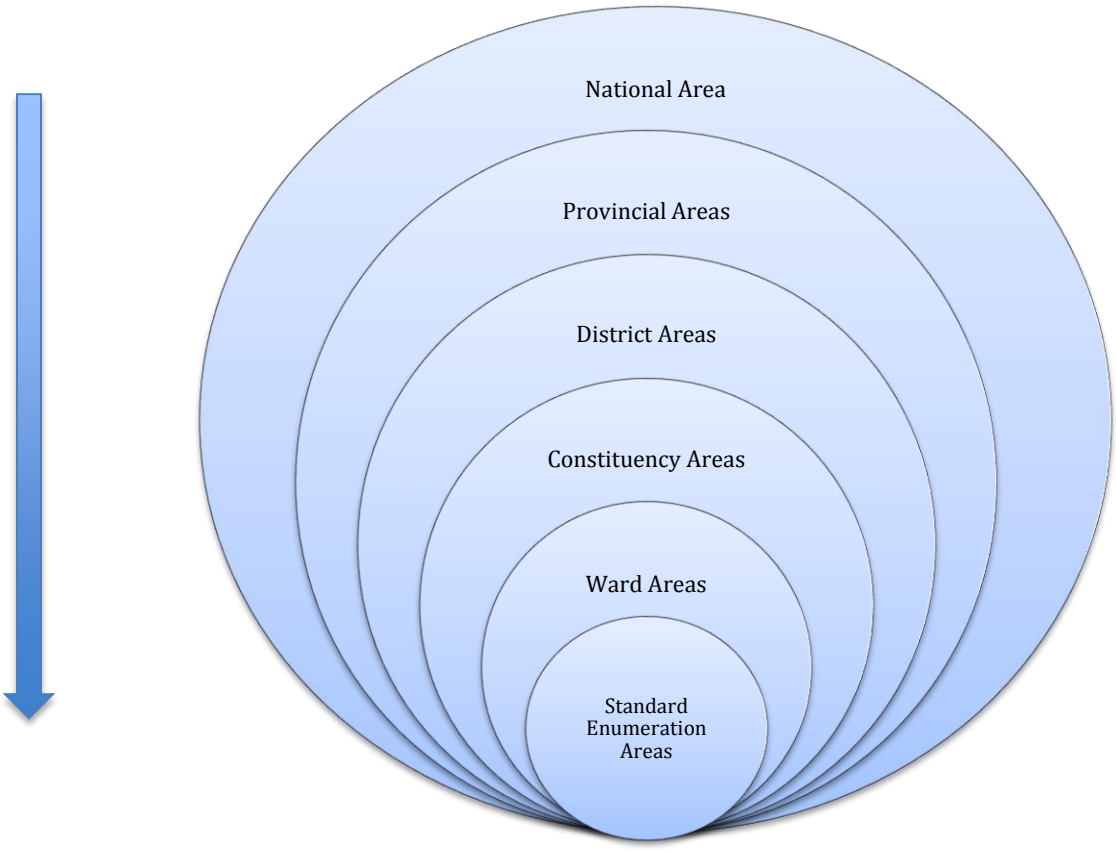


Figure 3.4 Zambia administrative boundary areas in descending hierarchal order

The spatial join created an attribute table that included all of the information from the GCC dataset for each boundary area and the spatial attributes from each administrative area boundary shapefile. The GCC dataset was now ready to preliminarily assess the spatial distribution of patient health outcomes (Figure 5.2).

3.1.3 CSO Population Datasets

Population data was also added to the dataset for each SEA administrative area prior to ArcGIS input. The Central Statistical Office provided shapefiles that were used to identify and manually enter population data for the village areas. The shapefile for the SEAs included the 2010 population estimates and the number of households for each SEA area. This information was spatially joined with the GCC attribute table. The SEA boundary population areas served as a proxy for CHW catchment areas. CHW catchment areas are artificially derived, unofficial geographic boundary areas that not available for mapping.

In addition to the SEA populations, the ward level populations were included as well. This population and administrative area was best for analyses regarding the RHC health outcomes. Each health facility operated in a catchment area that was at the ward level and would provide insights into the total number of cases per the total population in the area. This data was unfortunately unavailable electronically but was available as a hard copy booklet at CSO. A request for a copy of the booklet was granted and the population tables for Chongwe district and wards were used to identify the population distributions, both household and total population, for each ward area. This information was manually entered for the respective wards for each patient village location in the GCC dataset prior to upload into ArcGIS. There were 17 Ward areas within Chongwe district and both Ward level and SEA level population information were recorded manually into an Excel GCC dataset, which served as a backup to the ArcGIS dataset.

3.1.4 District Health Facility Dataset

In addition to the overall population distributions for the sub-district areas, the actual RHC catchment populations were also included in the GCC dataset. A request was made to the District Health Office for RHC catchment area populations. This information was then added to the GCC dataset for each RHC location, prior to the upload into ArcGIS.

3.2 Variables

The complete GCC spatial dataset included coordinate locations for villages and RHC, administrative area information, and health outcomes. The following table (Table 3.2) is an in-depth summary of each variable included in the GCC dataset and uploaded in ArcGIS after village location matching and census data were included. If the data methods had been performed by CHWs using a GPS enabled mHealth device for community health data management, this is an example of the type of data output that would be expected for use with a GIS.

Table 3.2 GCC Geodataset with census data and village locations identified using PATH database.

Variable Name	Description	Source
Provincial Code	This is a first order subnational administrative boundary area. The country of Zambia is divided into 10 provincial locations. The Central Statistics Office has assigned geoIDs to each provincial area available shapefiles. The geoID for Lusaka Province is 05 and is the ID used to spatially join the GCC dataset with provincial level CSO shapefiles.	CSO, 2010 Census
Provincial Name	This refers to the corresponding Provincial name for the geoID allocated to each provincial area. The geoID 05 is for the Lusaka Provincial administrative boundary area.	CSO, 2010 Census
District Code	This is a second order subnational administrative boundary area. At the time of the 2010 census, the country of Zambia was divided into 74 district level locations. The Central Statistics Office has assigned geoID to each district area in available shapefiles. The geoID for Chongwe district is 501 and is the ID used to spatially join the GCC dataset with district level CSO shapefiles.	CSO, 2010 Census
District Name	This refers to the corresponding District name for the geoID allocated to each district area. The geoID 501 is for the Chongwe district administrative boundary area.	CSO, 2010 Census

Constituency Code	This is a third order subnational administrative boundary area. At the time of the 2010 census, the country of Zambia was divided into 150 constituency area level locations. The Central Statistics Office has assigned geoIDs to each constituency area in available shapefiles. The geoIDs for the constituencies included in the GCC study, conducted in Chongwe district, are 501073 & 501074 and are the IDs used to spatially join the GCC dataset with constituency level CSO shapefiles.	CSO, 2010 Census
Constituency Name	This refers to the corresponding Constituency names for the geoIDs allocated to each constituency area in Chongwe district. The geoIDs 501073 & 501073 are for Chongwe and Rufunsa constituency administrative boundary areas respectively.	CSO, 2010 Census
Ward Code	This is a fourth order subnational administrative boundary area. At the time of the 2010 census, the country of Zambia was divided into 1,430 ward area level locations. The Central Statistics Office has assigned geoIDs to each ward area in available shapefiles. The geoIDs for the ward areas included in the GCC study conducted in Chongwe district, are 50107301, 50107302, 50107303, 50107304, 50107305, 50107306, 50107307, 50107308, 50107309, 50107310, 50107412, 50107413, 50107414, 50107415, 50107416, & 50107417 and are the IDs used to spatially join the GCC dataset with ward area level CSO shapefiles.	CSO, 2010 Census
Ward Name	This refers to the corresponding Ward names for the geoIDs allocated to each constituency area within Chongwe district. The geoIDs 50107301, 50107302, 50107303, 50107304, 50107305, 50107306, 50107307, 50107308, 50107309, 50107310, 50107412, 50107413, 50107414, 50107415, 50107416, & 50107417 are for Kapwayambale, Chinkuli, Ntandabale, Chongwe, Kanakantapa, Chalimbana, Nakatindi, Lukoshi, Manyika, Lwimba, Nyangwena, Bunda Bunda, Nyamanongo, Rufunsa, Mankanda & Shikabeta ward administrative boundary areas respectively.	CSO, 2010 Census

Ward Household Population	This refers to household composition at the ward area level, which classifies all households according to the relationships among the people in them, and whether there is a family nucleus present or not. These people may or may not be related by blood, marriage, or adoption but make common provision for food or other essentials for living and have one person whom is regarded as head. A household can also have one member.	CSO, 2010 Census
Ward Population	This refers to total population for Ward areas included in the GCC study. Population estimates were De Jure meaning that the usual household members present and usual household members temporarily absent at the time of census were counted as members of their usual households. The de jure counts is considered the most accurate population count for the country.	CSO, 2010 Census
SEA Code	These refer to geographic statistical unit numbers created in the Census of Population and Housing (CPH), which contain a certain number of households. SEAs are well-defined boundaries with area codes that are recorded on census maps. There were 334 SEA code areas for Chongwe district.	CSO, 2010 Census
SEA Population	An SEA is usually a group of small villages or a village, or a part of a large village in the rural areas. The SEA population is the total number of individuals living in an SEA area.	CSO, 2010 Census
SEA Household Population	An SEA is usually a group of small villages or a village, or a part of a large village in the rural areas. Household population is the number of households located within villages or number of families within the villages of an SEA area. .	CSO, 2010 Census

Date	The date the CHW saw the patient and health information was recorded into registry.	GCC CHW Registries, 2013
Rural Health Center Name	The name of the Rural Health Center that each CHW reported to in the GCC study within Chongwe district.	GCC CHW Registries, 2013
RHC Latitude	GPS latitude coordinates for Rural Health Center locations.	MOH 2012, CSO 2012
RHC Longitude	GPS longitude coordinates for Rural Health Center locations.	MOH 2012, CSO 2012
RHC Catchment Area Population	This refers the total community population that is served by a RHC facility.	Chongwe District Health Office Registries 2013
Patient Address	This is the village name of the patient that was recorded during the patient visit by the CHW into the registry.	GCC CHW Registries, 2013
Patient Address Latitude	This is the latitude coordinate for each village area identified using CSO census shapefiles and PATH program registries.	CSO Census 2010, PATH
Patient Address Longitude	This is the latitude coordinate for each village area identified using CSO census shapefiles and PATH program registries.	CSO Census 2010, PATH
Male	Patients identified as 'Male' gender in CHW registries	GCC CHW Registries, 2013
Female	Patients identified as 'Female' gender in CHW registries	GCC CHW Registries, 2013
Age	Patients Age	GCC CHW Registries, 2013

Fever	Patients were diagnosed with a fever by CHW during visit	GCC CHW Registries, 2013
History of Fever	Patients who had a fever in the past month	GCC CHW Registries, 2013
Headache	Patients who complained of a headache during CHW visit	GCC CHW Registries, 2013
Cough	Patients who complained of a cough during CHW visit	GCC CHW Registries, 2013
Diarrhea	Patients who complained of diarrhea or watery stool during CHW visit	GCC CHW Registries, 2013
Vomiting	Patients who complained of vomiting during CHW visit	GCC CHW Registries, 2013
Problems Breathing	Patients who complained of problems breathing during CHW visit	GCC CHW Registries, 2013
Chest pain	Patients who complained of chest pain during CHW visit	GCC CHW Registries, 2013
No Symptoms	Patients who did not have or complain of any symptoms during CHW visit	GCC CHW Registries, 2013
Other Symptoms	Patients who complained of other symptoms during the CHW visit	GCC CHW Registries, 2013
ITNS	Patients who had Insecticide Treated Bed-nets	GCC CHW Registries, 2013
Slept Under ITNS	Patients who slept under an ITN the previous night	GCC CHW Registries, 2013
IRS	Patients who had Indoor Residual Sprays done in the home in the last 6 months	GCC CHW Registries, 2013
IPT	Intermittent Preventive Therapy for malaria provided to pregnant women using Sulfadoxine and Pyrimethamine drug therapies.	GCC CHW Registries, 2013

Positive Rapid Diagnostic Test	Patients who were given a RDT resulting in a positive confirmation for malaria	GCC CHW Registries, 2013
Negative Rapid Diagnostic Test	Patients who were given a RDT resulting in a negative confirmation for malaria	GCC CHW Registries, 2013
Rapid Diagnostic Test Not Done	Patients who were not given a RDT to confirm malaria	GCC CHW Registries, 2013
Coartem	This is the first line ACT drug treatment for malaria in Zambia	GCC CHW Registries, 2013
Fansidar	This is the first line IPT drug treatment for malaria in Zambia, it is provided as prophylaxis against malaria for pregnant women.	GCC CHW Registries, 2013
Amoxyl	The antibiotic Amoxicillin prescribed for infections.	GCC CHW Registries, 2013
Zinc	Zinc supplements provided for dehydration related to diarrhea. Often combined with ORS for rehydration therapy.	GCC CHW Registries, 2013
ORS	Oral Rehydration Salts (ORS) used to treat severe dehydration related to diarrhea in combination with Zinc supplements.	GCC CHW Registries, 2013
Panedol	Common pain killer that is prescribed directly by CHW to patients in the community	GCC CHW Registries, 2013
Other Drugs	Other drugs that are prescribed for community illnesses	GCC CHW Registries, 2013
Referred	Patients who were referred to nearest RHC for drug prescriptions and further follow-up.	GCC CHW Registries, 2013

3.3 Analysis

3.3.1 ArcGIS Data Preparation

Once all variables from each CHW registry were incorporated into the GCC geodataset, the dataset was imported into SPSS 15.0 statistical analysis software to perform descriptive statistical analysis. Cross-tabulations and frequencies were performed for each disease and illness outcome of interests for each village location area. The total number of fever, fever and malaria, malaria, fever with cough and problems breathing, fever and diarrhea, fever and vomiting, pneumonia, and severe diarrhea cases were tabulated for each village location. This information was added to a secondary geodataset in an Excel spreadsheet for each village location. Afterwards, this geodataset was imported into ArcGIS 10.2 and spatially joined to the existing GCC shapefile. Prior to adding any data into ArcGIS, the data frame was set to a projected coordinate of Arc 1950 UTM Zone 36S. All shapefiles, including village and RHC coordinate points were set to the same projection in preparation for analysis and mapping.

Finally, to determine the corresponding SEA for each village coordinate location, a shapefile of SEA areas was used and village coordinates were overlaid onto the SEA shapefile. A manual search was done using the 'identify tool' in ArcMap for each village location point placement to identify the corresponding SEA in which each village point fell. For each village location that fell within an SEA area, the SEA number was identified and recorded into a new field in a third dataset for each village location in the attribute table. After all SEA numbers were found for each of the patients and village locations, the dataset was spatially joined to the GCC shapefile. To disease distribution at the Ward level, a spatial join was also done using the Ward codes and Ward area level shapefiles. Information on corresponding village locations was available in the CSO report and was also recorded during the same time population estimates were determined for the Wards.

3.3.2 Population Density Distributions

To determine the population densities for both the Ward and SEA areas, the area of the various geographic boundaries were calculated first. The geographic areas were determined by creating a new field for each SEA and Ward level geography areas in ArcMap. In ArcMap the new field, 'Area', was added to the attribute tables and areas were calculated in square kilometers for each location and administrative level using the 'Calculate Geometry' function. After areas were determined, the population densities were calculated into an additional field labeled, 'Pop_desit', in the attribute tables using the 'Field Calculator' function to divide the total populations by the total square kilometer areas. The result was used to map the population densities for both of the respective SEA and Ward area levels. No population counts were missing and a few areas had a total population estimate of zero. Chloropleth maps showing the population densities using quintiles were created to evaluate the population distributions and case count distributions.

3.3.3 Illness Case Count Distributions

Illnesses were cross-tabulated with patient addresses to determine the locations and case counts using SPSS 15.0. More complex illnesses were cross-tabulated concordantly with two or more symptoms to determine these illness outcomes. For example, the illnesses '*Severe diarrhea'* and *'Pneumonia'* were defined as cases where both vomiting and diarrhea were present and where cough and problems breathing were present, respectively. Malaria cases were defined as occurrences where RDT results were positive and where RDT results were positive and fever was also present. Other febrile illness cases were tabulated and recorded as the total number of cases for each location into geodataset. Cases were determined for both the SEA and Ward area levels. Maps of proportional case counts for each Ward level were produce for exploratory data analysis and descriptive epidemiology of the general distribution of illnesses.

3.3.4 Incidence Rates

Once the total numbers of case counts were determined for each location, the incidence rates were calculated in Excel using the formula *cases/total population * 10^2* annually. 10^2 was used as the per person number at risk value due to the small areas and low population distributions associated with rural districts. The incidence rate for each location was calculated for all symptoms and illnesses of interest and recorded into the geodataset. Illness incidence rates were calculated for SEA administrative area level only to provide better spatial insights into symptom and illness distributions, at the community level. Chloropleth maps showing the incidence rates per 100 persons using natural breaks for descriptive spatial epidemiology. The incidence maps created and opportunity for understanding of the spatial distribution of diseases and for comparison with future spatial analysis results.

3.3.5 SaTScan v9.3 Cluster Analysis

SaTScan v9.3 Data Preparation

SaTScan v9.3 requires multiple input datasets to analyze spatial data for clustering or hot spots. Dataset inputs for SaTScan v9.3 analysis were prepared using Excel, ArcGIS 10.2, and SPSS 15.0. SaTScan v9.3 datasets consisted of a case file, population file, coordinate file, and an adjustment file for each disease/illness evaluated. All village level patient cases were aggregated using SPSS 15.0 to the SEA area level for each illness case file as total counts.

The total population for each SEA area was used for the population file. Missing population data or zero population areas were identified and entered into a separate adjustment file used for spatial adjustments made by the software during analysis. Geographic coordinate locations for each SEA were required for the coordinate file input for SaTScan v9.3 analysis. Coordinate locations for each SEA area were determined using ArcGIS to identify the centroid location of each polygon feature in the SEA shapefile. The resulting output provided the latitude

and longitude coordinates for each SEA location. This file was extracted into Excel and a 'vlookup' formula using the SEA codes as the matching variable was used to match the geodataset with all the SEAs coordinate locations. This would allow for identification of SEA areas identified to be cluster areas in the SaTScan v9.3 analysis using ArcGIS.

Missing data for locations and cases were included in the adjustment file for areas with no data. The adjustment file also included risk estimates for areas with missing population information for every disease and illness. A relative risk of 0 was determined for these locations in the adjustment file, used during analysis.

SaTScan v9.3 Methods

SaTScan v9.3 was used to run a standard purely spatial scan statistic for each illness and disease outcome of interest. Spatial cluster analyses were performed to test whether disease/illness cases were distributed randomly over space and, if not, to identify areas of spatial disease clusters for statistical significance. Spatial scan statistics for cluster detection was applied to test the null hypothesis that the relative risk (RR) for disease/illness cases were the same between SEAs or groups of SEAs.

The SaTScan v9.3 methods imposed a circular window on the map, which moved over the areas and centered on the centroid of each SEA, allowing the maximum window for cluster size to be set to any value less than 50% of the total population. The default setting was 50% of the total population at risk and this was maintained for all spatial analyses conducted (Kulldorff 2014). This allowed SaTScan v9.3 to evaluate very small and very large clusters, and everything in between. For each window of varying population and size, the software tested the risk of illness inside and outside the window. Using the null hypothesis of spatial randomness, the expected number of cases in each window was proportional to the combined population of SEAs whose centroid is inside the circle. This allowed for SaTScan v9.3 adjusted for uneven

population distributions. The datasets were scanned for only clusters with high rates of disease/illness, equivalent to a one-sided statistical test (Kulldorff 2014).

Clusters were identified by a comparison of the expected and observed number of cases within and outside a scanning window that has a varying radius and center. This comparison is called the Likelihood Ratio (LR), and it determined how likely a cluster exists due to chance. The higher the LR, the more likely the cluster exists or the maximum likelihood, due to more than chance alone. The scanning window with the maximum likelihood ratio is flagged as the most likely or primary cluster and LR is reported. The primary cluster and significant secondary clusters (ordered in descending LR values) are also included in the output. P-values were assigned to each cluster to show statistical significance of cluster findings, with small p-values indicating clusters of significance. The p-values were obtained through Monte-Carlo simulations that randomly generated 999 replications of the dataset under the null hypothesis. Monte-Carlo simulations create random simulations of the data, then report how many of those simulations resulted in higher likelihood ratios than what the actual data found (Kulldorff 2014).

Clusters that were identified for each disease or illness of interest were mapped in ArcGIS to visualize the location of significant primary and secondary clusters. Areas identified to be clusters area were shaded in pink for secondary clusters and a darker shade of red for primary clusters for each disease/illness outcome evaluated. Clusters with p-values less than .05 were considered significantly valid clusters. The SaTScan v9.3 scanning window buffer areas for each cluster group was also included in the maps at the 50% risk population level.

A relative risk and its p-value were also given for each cluster output, indicating the risk of developing a disease/illness within cluster or geographic areas relative to areas outside of those locations. In epidemiology, a relative risk of 1 would indicate equal risk for disease incidence given exposure. A relative risk above or below 1 would indicate an elevated or

decreased risk, respectively, for disease incidence given exposure. Relative risks that are below 1 are said to have a protective effect, meaning that exposure may prevent or delay disease or illness incidence. In SaTScan v9.3, a relative risk above 1 in a cluster indicated how much more common a disease/illness occurred in those locations compared to the baseline (Kulldorff 2014). The corresponding p-values for relative risks determine the probability that these could have happened by chance alone.

CHAPTER FOUR: RESULTS

4.1 Spatial Epidemiologic Maps

4.1.1 Population Density Maps (Ward and SEA)

The population density of the Ward area level was mapped to visualize the population distribution throughout the district (Figure 4.1). Darker areas of blue indicate higher population densities for the Ward areas. The distribution of village and Rural Health Center locations can also be observed. The Ward area level maps provide insights at the sub-district level for Rural Health Centers. Rural Health Centers in the Western part of the district are located in wards with the highest population densities. Rural Health Centers in the Eastern part of the district were less in number as a result of lower population densities. The village locations were spread throughout the district and were greater in the Western part of the district and along the major roadway, the T4, moving from West to East away from Lusaka city.

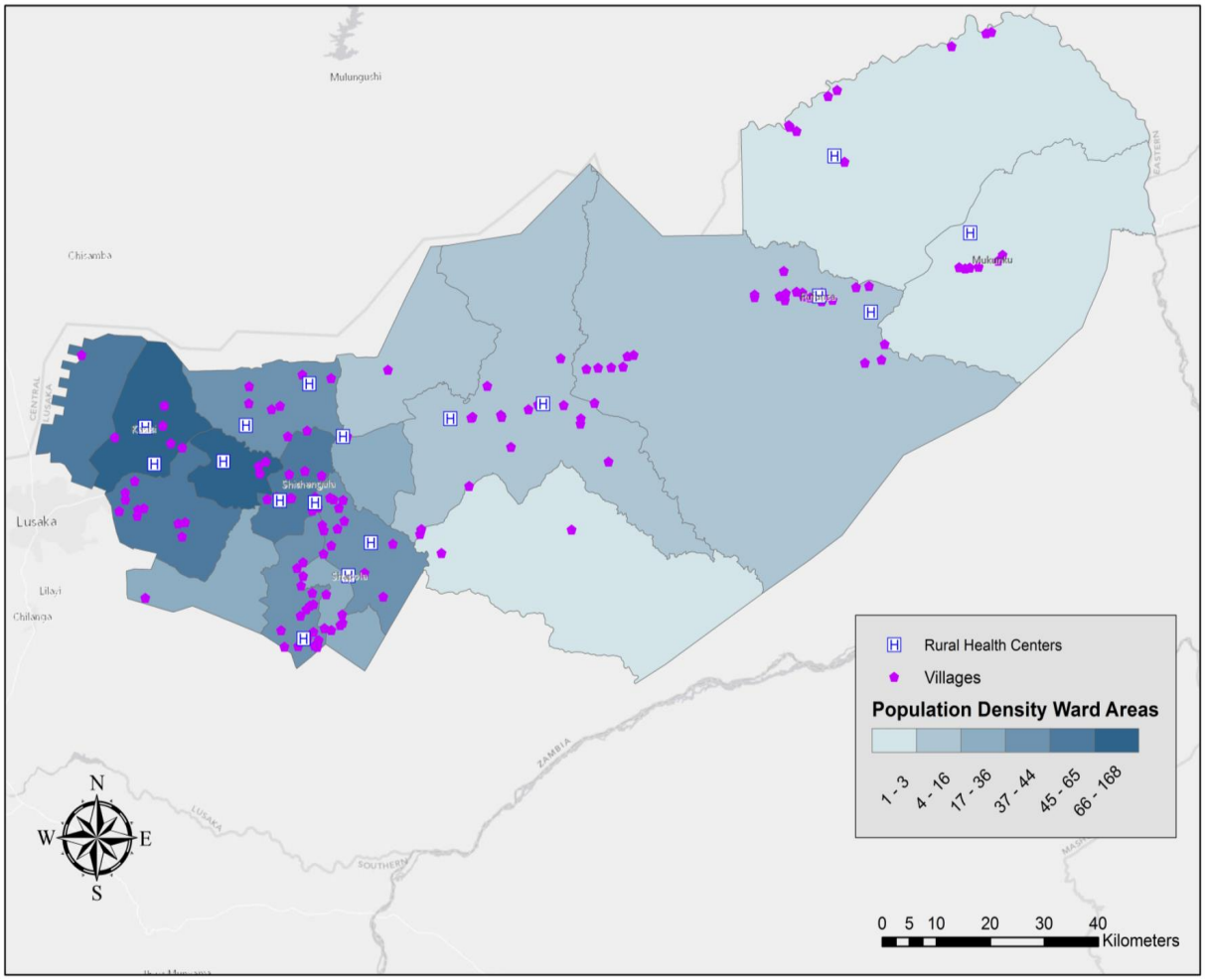


Figure 4.1 Chongwe district population densities for Ward administrative areas with village and RHC locations

The population density for Chongwe SEAs was also mapped to evaluate the population distributions at the community sub-district level.

The population distribution at the SEA level was similar to the Ward areas, however the population was most dense in the center of the Western section of the district area (Figure 4.2.). Village locations are represented in pink dots were overlaid onto a population density map. The most densely populated SEA areas also had the smallest land area in the district. The population densities and numbers of RHC locations also coincided at this administrative area level, with areas with higher densities having more RHC facilities than lower densely populated areas.

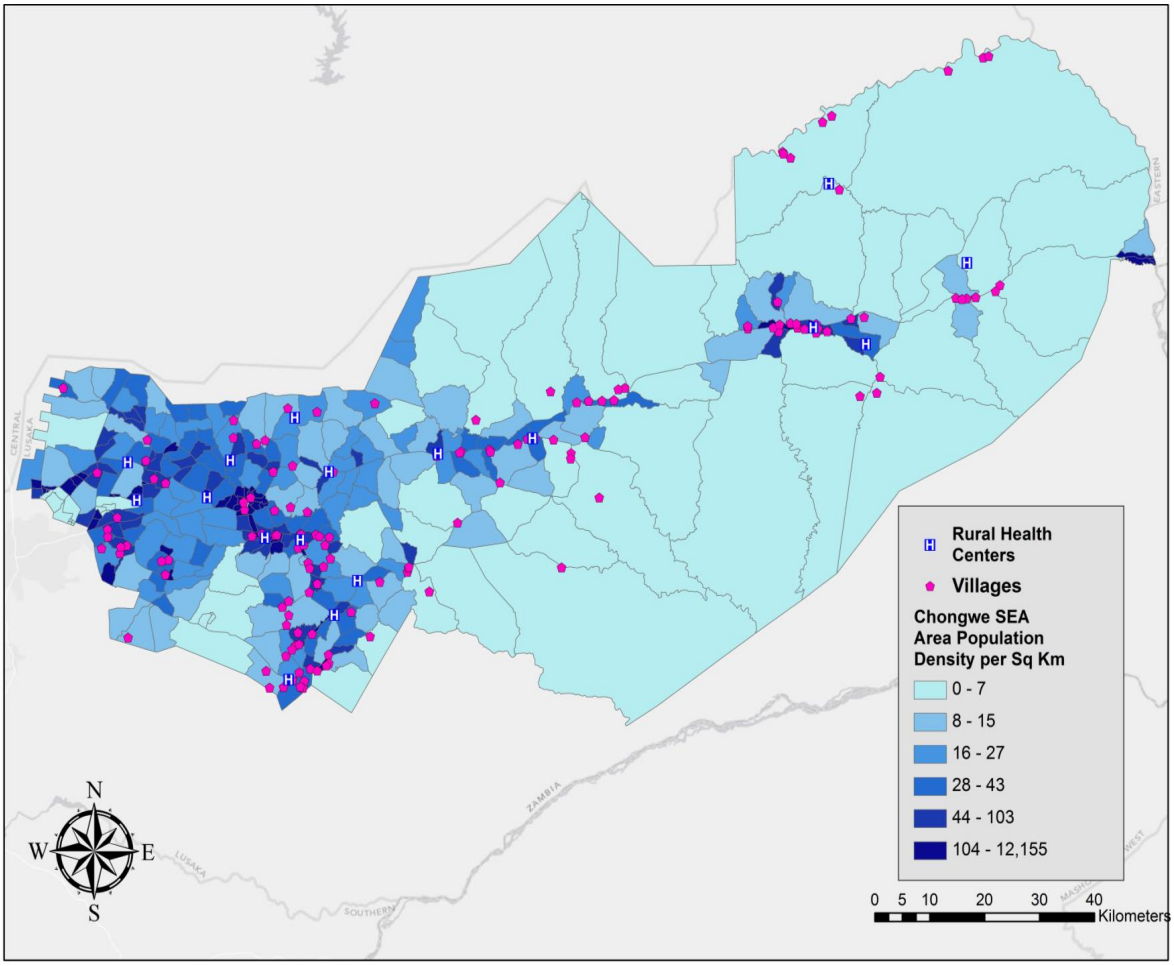


Figure 4.2 Chongwe population densities for SEA administrative boundary areas and RHC locations

4.1.2 Illness Case Count Distributions

Case counts were determined for each disease/illness at the Ward and SEA area level. The counts were shown using proportional symbols to illustrate the volume of cases for each ward area in comparison to other ward areas. These proportional symbol maps helped to visually evaluate different disease/illness distributions for various locations. Larger circles indicate more cases or counts and smaller ones proportionately fewer counts. For some maps, multiple variables were overlaid to identify areas were symptoms occurred together or to evaluate number cases versus interventions. Three proportional symbol variables were overlaid together on the

same map (Figure 4.3); malaria, fever, and RDT usage to identify areas where fever counts were high and malaria cases were low, and where fever cases were high and RDT counts use was low to see possible areas of concern for effective case management. Areas were identified where RDT use was proportionately less than in areas with high fever and malaria counts, indicating a need for more RDT implementation in these areas.

The proportions of cases in the Wards indicated areas where possible clusters of illness may exists and were potential areas for further interventions. There were areas where malaria cases were low, but fevers were high and RDT usage was also high to confirm truly confirm malaria (Figure 4.4). These areas would be priority areas to explore further for other causes of fevers and to develop interventions to address these unknown causes of fever. Additionally, there were areas (Eastern part of district) where malaria cases were higher in proportion to RDT usage, and these areas are of equal concern because they may be potential areas for targeted RDT resource allocation.

4.1.3 Incidence Rates

The SEA administrative level was used for community level calculations and all geostatistical analyses of malaria, pneumonia, and severe diarrhea illness outcomes and to map the incidence rates for all diseases/illnesses. This administrative level was the closest proxy to CHW catchment areas and community level assessments. Incidence rates were calculated for each illness recorded in the CHW registries and were rendered in ArcGIS in chloropleth maps (Figure 4.4).

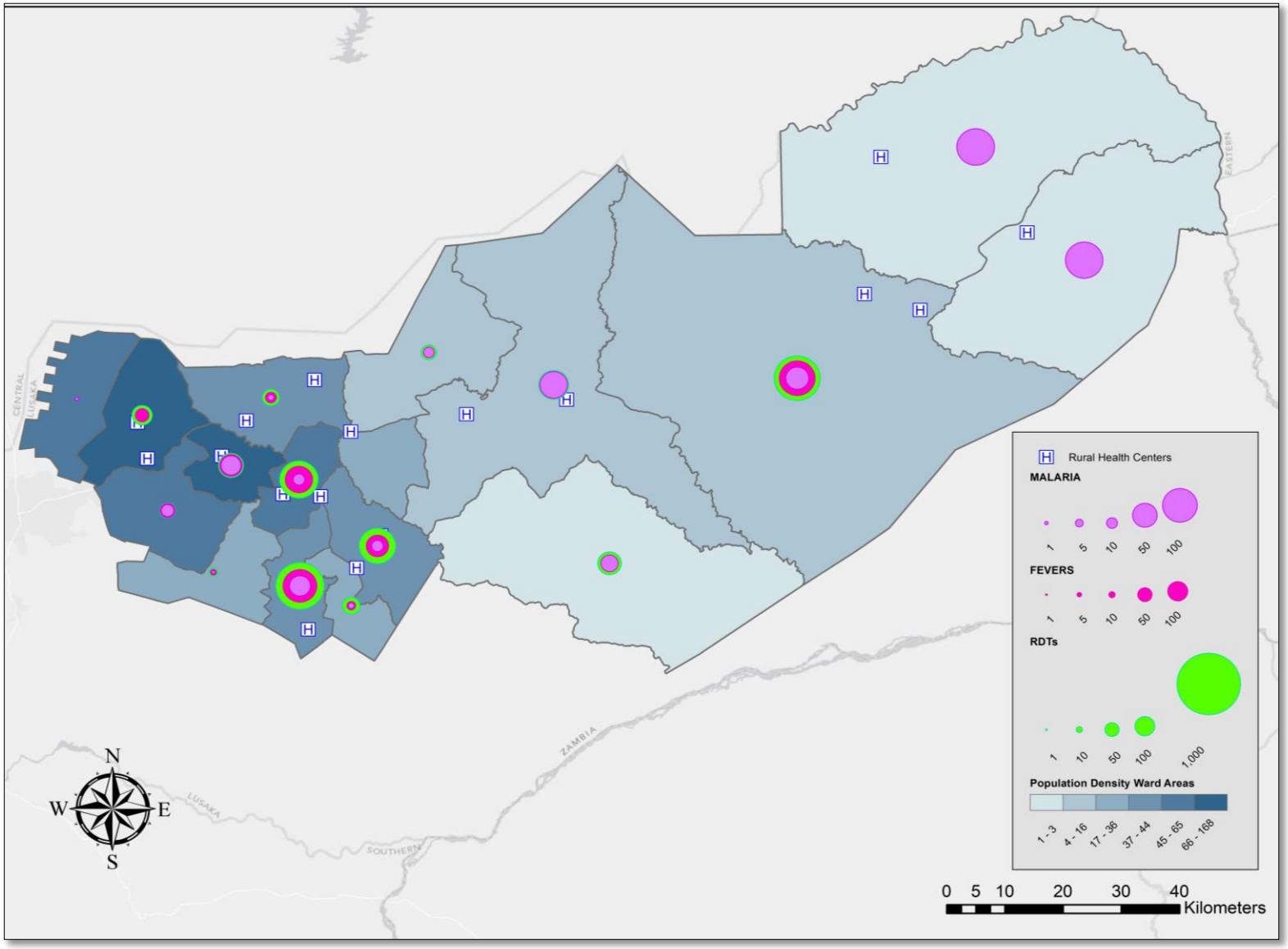


Figure 4.3 Proportional Symbol Map of case counts for malaria, fevers, and RDT usage within Chongwe Ward areas.

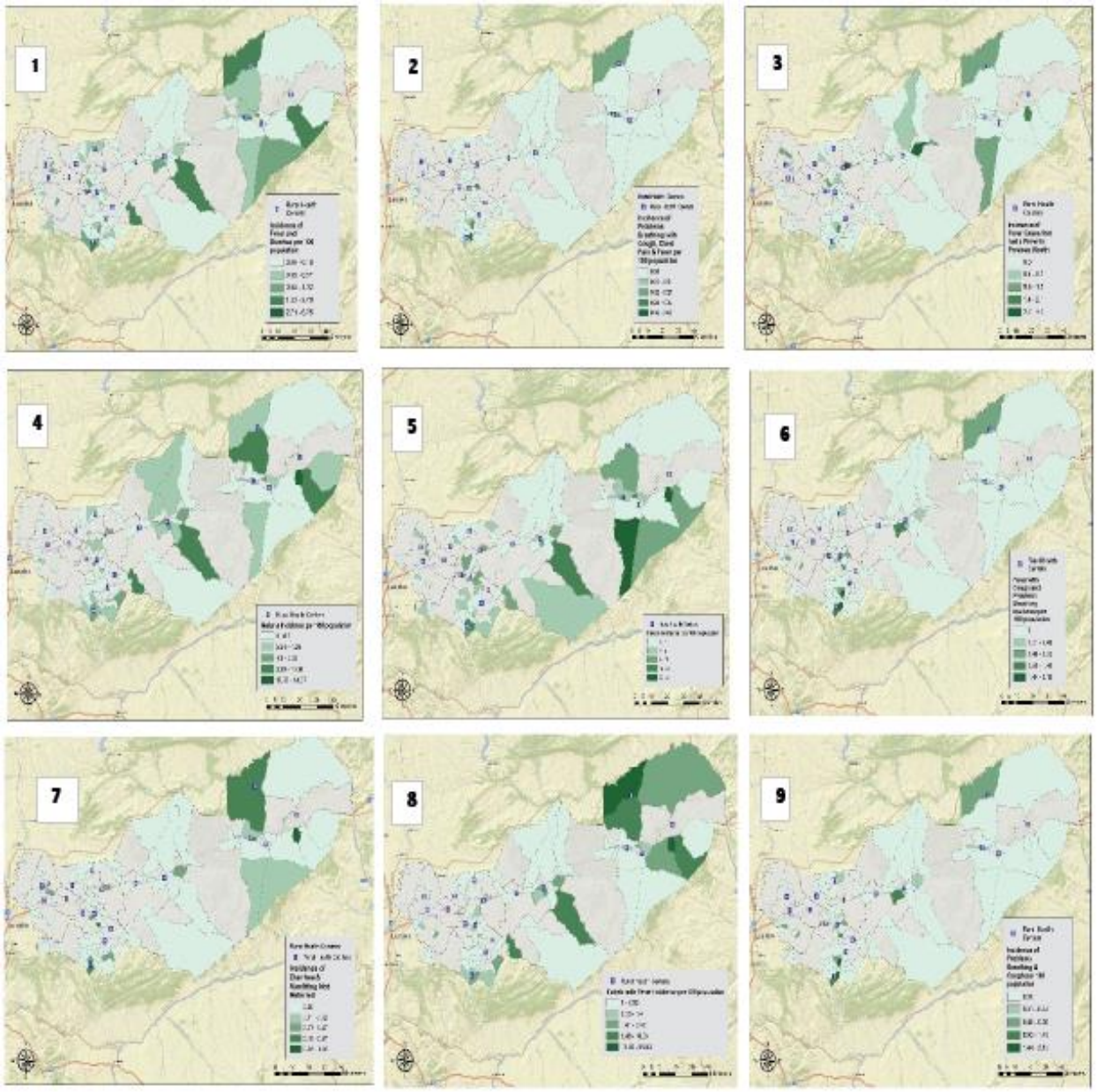


Figure 4.4 Chongwe disease/illness incident rate estimates per 100 persons annually for SEA administrative boundary areas: 1.) Incidence of Diarrhea & Fever; 2.)Incidence of Problems Breathing, Chest Pain, Cough and Fever; 3.) Incidence of Fevers with Fever History; 4.) Malaria Incidence; 5.) Incidence of Fever; 6.) Incidence of Fever, Cough, & Problem Breathing; 7.) Incidence of Severe Diarrhea (Diarrhea & Vomiting) not referred; 8.) Incidence of Malaria with Fever; 9.) Incidence of Cough & Problems Breathing (Pneumonia)

An evaluation of the incidence maps provided additional insights into the spatial distribution of disease/illness cases. There were variations in incidence rates for malaria, severe diarrhea, and pneumonia throughout the district. For a number of diseases/illness, the highest

incidence rates occurred in remote and less densely populated areas (Figure 4.4. Map #: 1,4,5,7,8). For malaria incidence, areas known to have the greatest risk for malaria outcomes coincided with map areas that had the highest incidence rates.

4.2 Cluster Analysis Results

Sub-district areas identified on the maps to have the highest incidence rates were locations that required further investigation. A spatial scan cluster analysis was performed to provide more in-depth insights into these high incidence areas. The analysis focused on fever, malaria, diarrhea, vomiting, severe diarrhea, and pneumonia disease/illness burdens, which were identified by CHWs to be the most common in the district. Areas identified to be statistically significant clusters were represented with pink shading. Primary cluster areas where represented with a darker shade of pink in all maps. Secondary clusters continued to be represented with a lighter shade of pink for all maps.

If areas of high disease/illness incidence were identified to be statistically significant hot spots, then it would provide evidence for stakeholders and CHW to make better-informed decisions inform where allocation of scarce resources would be most effective for treatment and case management by CHWs.

4.2.1 Fever

The cluster analysis of fever incidence identified seven primary and secondary cluster areas (Figure 4.5). The scan identified seven cluster areas. The primary most likely cluster area consisted of 15 rural SEA areas near Rufunsa RHC and Mpanshya Mission Hospital RHC, indicated in darker pink in the figure. This group of SEAs was determined to be the most likely area to have a cluster of fever incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for the primary cluster area was 8.97 ($p < 0.05$), indicating that the risk of fever within this area group was greater than locations outside this area group. There was one SEA area, shown in the

inset map, located in a high population density area that had an excess risk of fever incidence (RR= 41.2; p>.05). This area was determined to be a secondary cluster area.

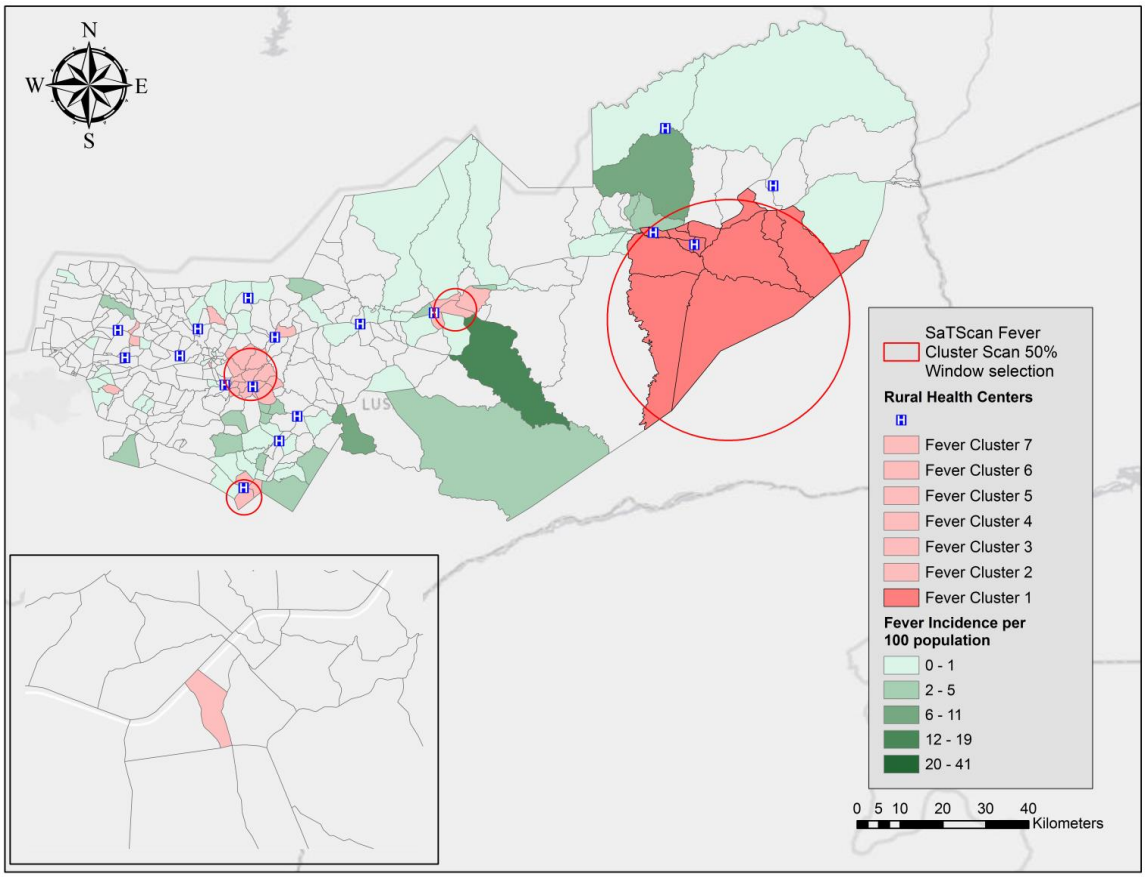


Figure 4.5 Fever Incidence Cluster Areas

4.2.2 Malaria

The cluster analysis of malaria incidence identified five primary and secondary cluster areas. The scan identified a primary cluster consisting of 2 rural SEA areas near Lukwipa RHC (Figure 4.6). This highlighted group in the darker shade of red was the group of SEAs was determined to be the most likely area to have a cluster of malaria incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for the primary cluster area was 57.4 ($p < 0.05$), indicating that the risk of malaria within this area group was more than locations outside this area group. There were two SEA areas, located in a high population density areas that had an excess risk of

fever incidence (RR= 58.2; p>.05; RR= 44.0; p>.05), however there were determined to be a secondary cluster areas and were not identified to be a most likely cluster area. Although the RR were high in these areas, it was not unusual given the malaria endemicity and spatial variability in malaria outcomes.

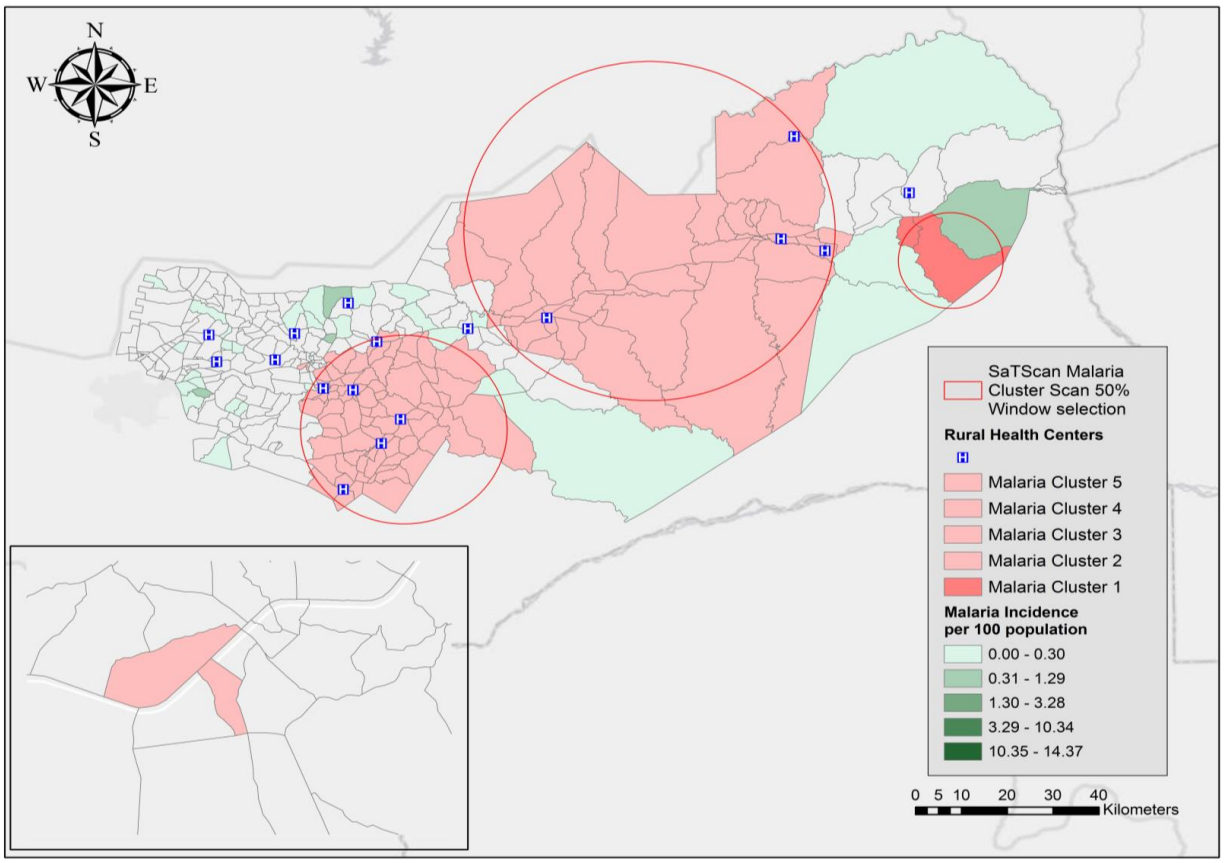


Figure 4.6 Malaria Incidence Cluster Areas

A secondary analysis of recorded malaria cases was done, but with the inclusion of the presence of a fever symptom (Figure 4.7). This analysis of the incidence of malaria outcomes with fever identified six primary and secondary cluster areas. The scan identified a primary cluster consisting of 23 rural SEA areas near Lukwipa RHC, Shikabeta RHC, Rufunsa RHC, and Mpanshya Mission Hospital RHC in the Eastern part of the district and Kasisi RHC and Waterfalls RHC in the Western part of the district. In addition to this cluster group, four SEA

areas in the more densely populated Western part of the district were also included as primary cluster areas. This group and location of SEAs was determined to be the most likely areas to have a cluster of malaria with fever incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for the primary cluster areas was 20.3 ($p < 0.05$), indicating that the risk of malaria with fever within these group and locations was more than locations outside this area group and locations.

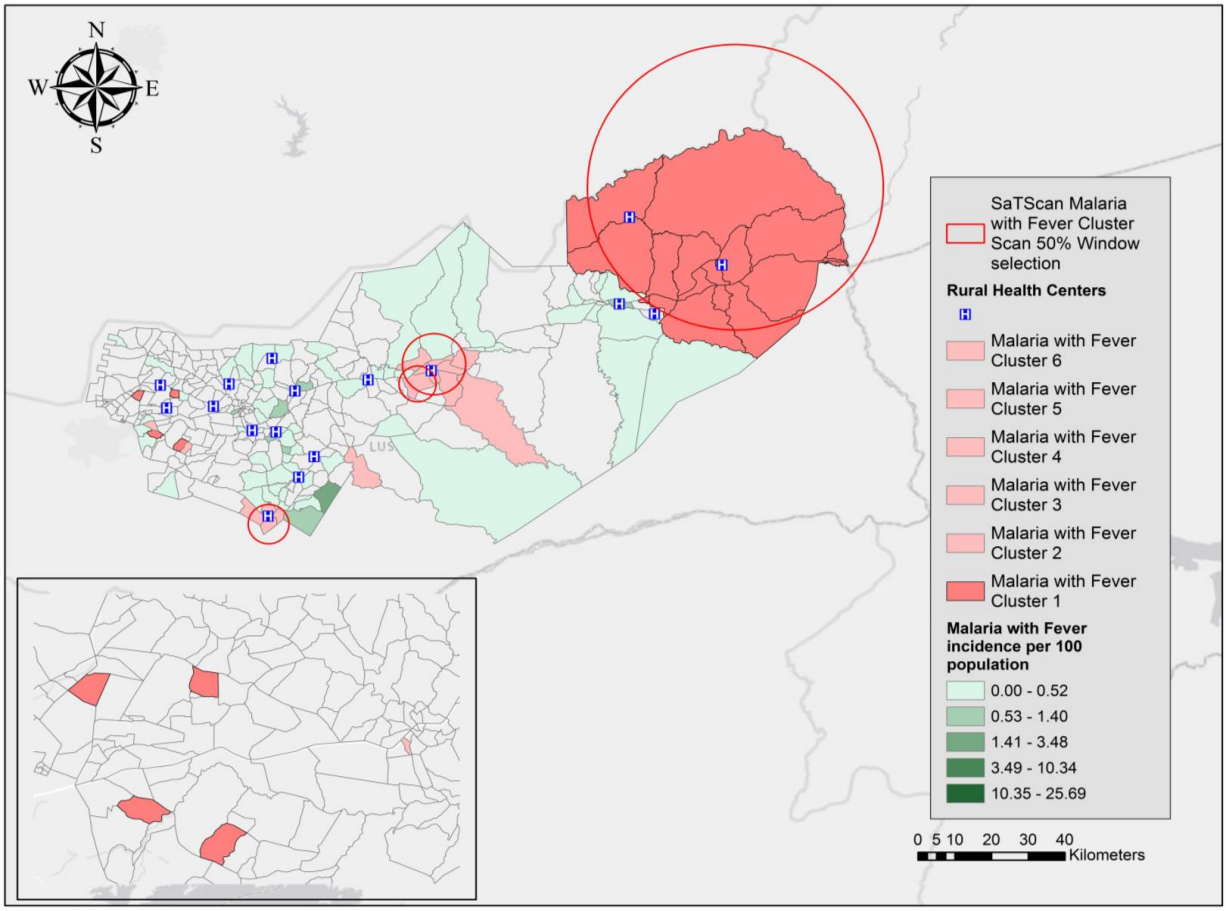


Figure 4.7 Malaria with Fever Incidence Cluster Areas

4.2.3 Diarrheal Illnesses

The case definitions for diarrheal illnesses were vomiting, diarrhea, and for severe diarrhea outcomes, the presence of both diarrhea *and* vomiting symptoms together. The cluster analyses evaluated each of these diarrheal illness outcomes to identify sub-district level hot spot areas in Chongwe district (Figures 4.8 and 4.9).

4.2.3.1 Diarrhea

The cluster analysis of diarrhea illness incidence identified three primary and secondary cluster areas. The scan identified a primary cluster consisting of 39 SEA areas within a densely populated area in the Western part of the district, near Katoba RHC, Lwiimba RHC, Mwalumina RHC, Chalimbana RHC, and Kampekete RHC. This group of SEAs was determined to be the most likely area to have a cluster of diarrhea incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for the primary cluster area was 4.76 ($p < 0.05$), indicating that the risk of diarrhea incidence within this area group was more than locations outside this area group.

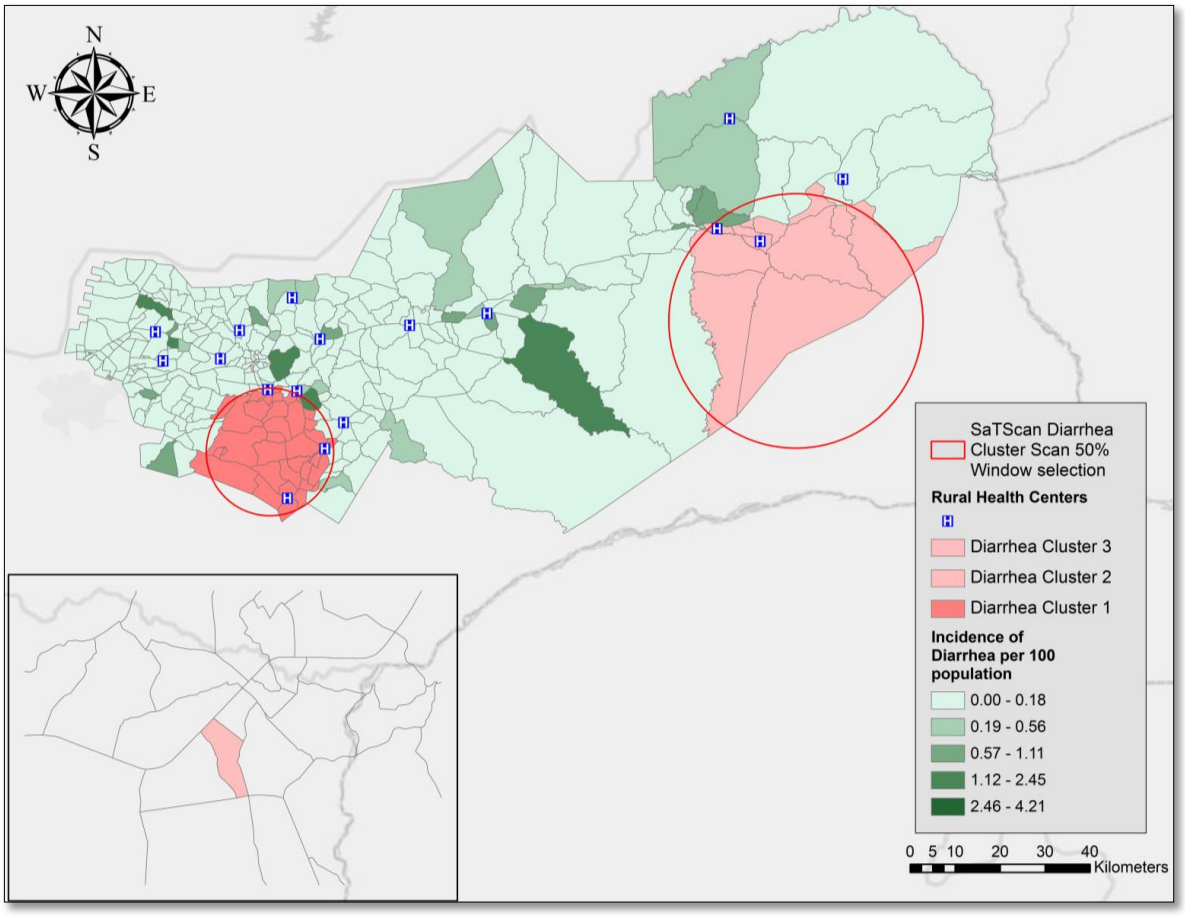


Figure 4.8 Diarrhea Incidence Cluster Areas

4.2.3.2 Vomiting

The cluster analysis of only vomiting illness incidence identified five primary and secondary cluster areas (Figure 4.9). The scan identified a primary cluster in a single SEA area within a densely populated area in the Western part of the district near Kasisi RHC and Waterfalls RHC. This SEA area, which is shown in the inset map, was determined to be the most likely area to have a cluster vomit incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for this primary cluster area was 51.71 ($p < 0.05$), indicating that the risk of vomiting illness incidence within this SEA area was more than any other location outside of this area.

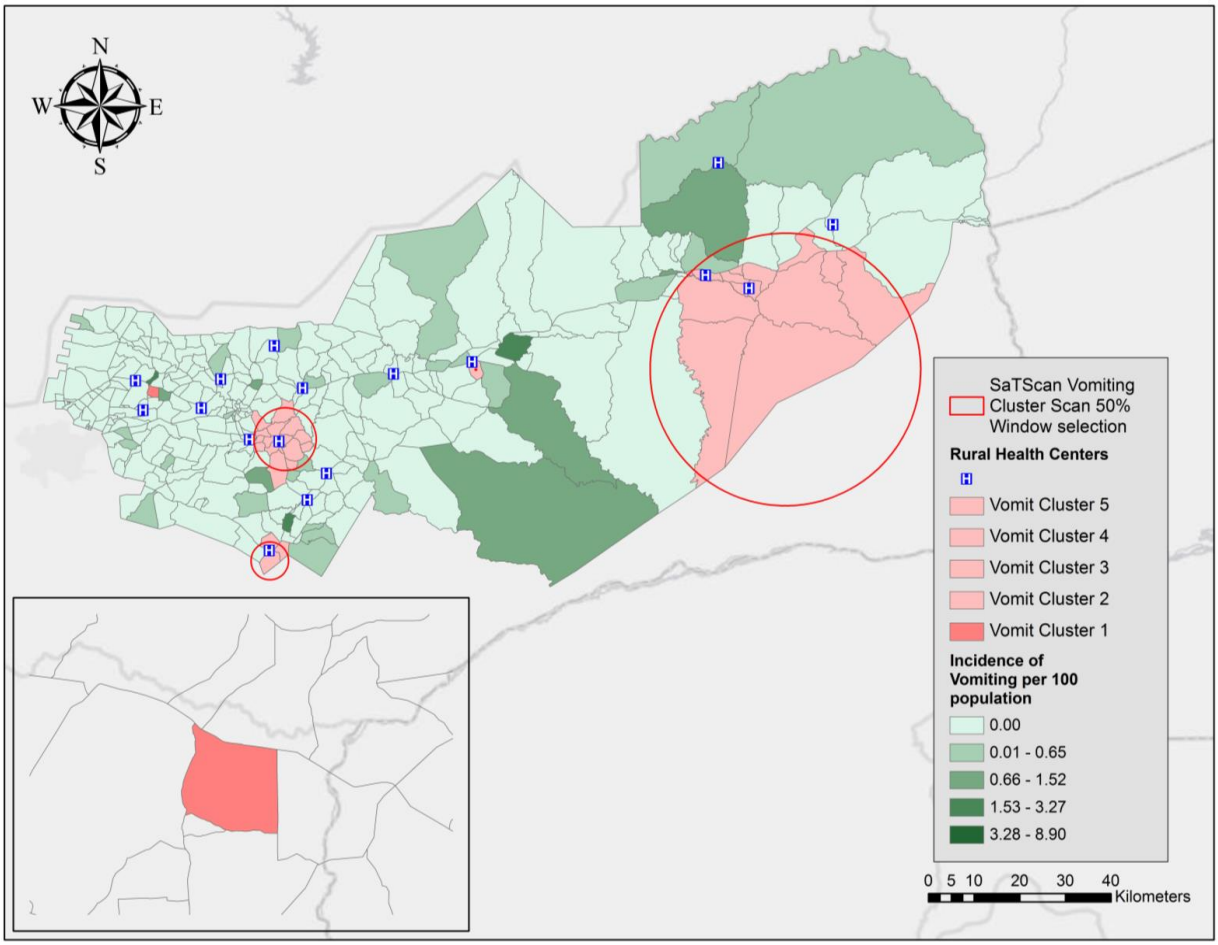


Figure 4.9 Vomiting Incidence Cluster Areas

4.2.3.3 Severe Diarrhea

The cluster analysis of severe diarrhea illness incidence identified four primary and secondary cluster areas (Figure 4.10). The scan identified a primary cluster consisting of 11 SEA areas within a densely populated area in the Western part of the district, near Chalimbana RHC and Kampekete RHC. This group of SEAs was determined to be the most likely area to have a cluster severe diarrhea incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for the primary cluster area was 10.61 ($p < 0.05$), indicating that the risk of severe diarrhea incidence within this area group was more than locations outside this area group.

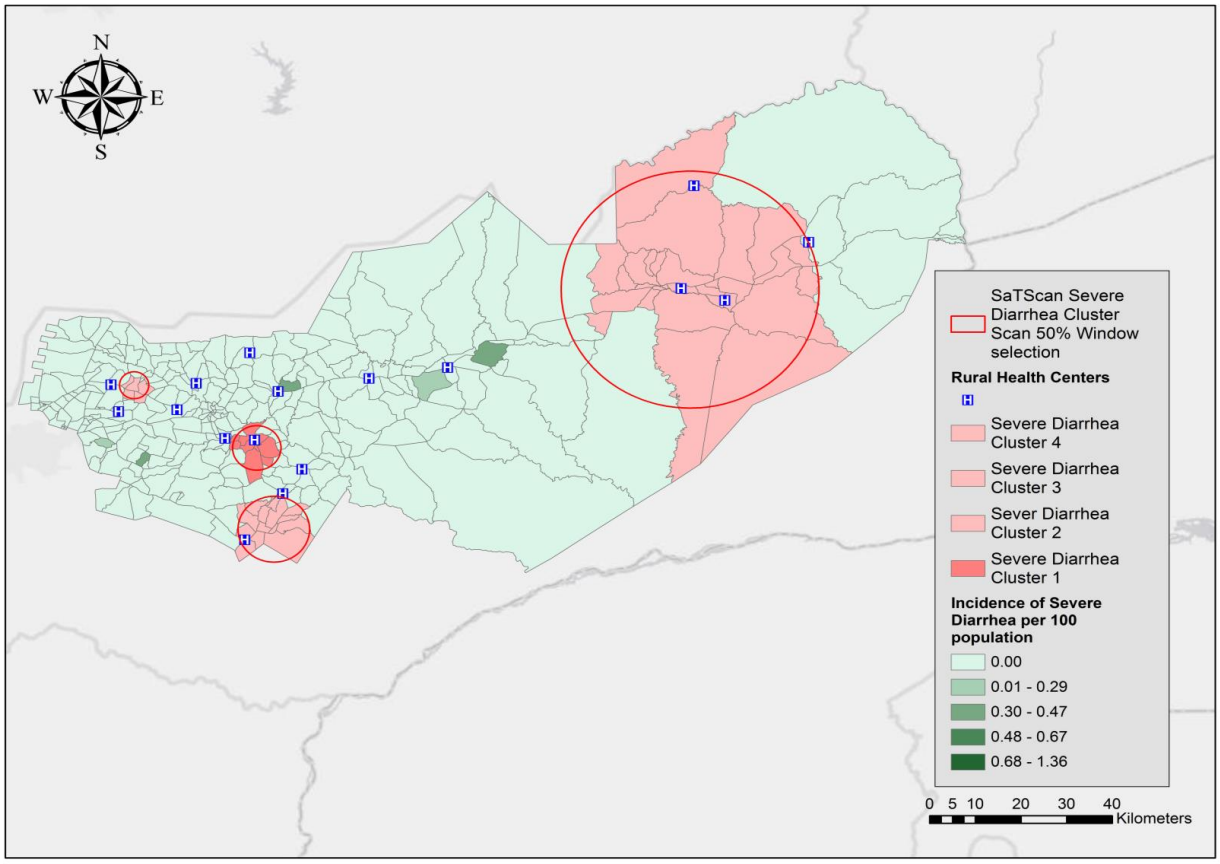


Figure 4.10 Severe Diarrhea Incidence Cluster Areas

4.2.4 Pneumonia

The case definition for pneumonia illnesses were cough *and* problems breathing and for severe pneumonia, the definition was the presence of fever, cough *and* problems breathing symptoms occurring together. The cluster analyses evaluated both pneumonia and severe pneumonia illness outcomes to identify sub-district level hot spot areas in Chongwe district.

4.2.4.1 Pneumonia

The cluster analysis of pneumonia incidence identified three primary and secondary cluster areas (Figure 4.11). The scan identified a primary cluster consisting of 10 SEA areas in the Western part of the district, near Katoba RHC. This group of SEAs was determined to be the most likely area to have a cluster of diarrhea incidence with a Monte Carlo rank of 1/1000 ($p < 0.05$). The RR for the primary cluster area was 23.75 ($p < 0.05$), indicating that the risk of pneumonia incidence within this area group was more than locations outside this area group.

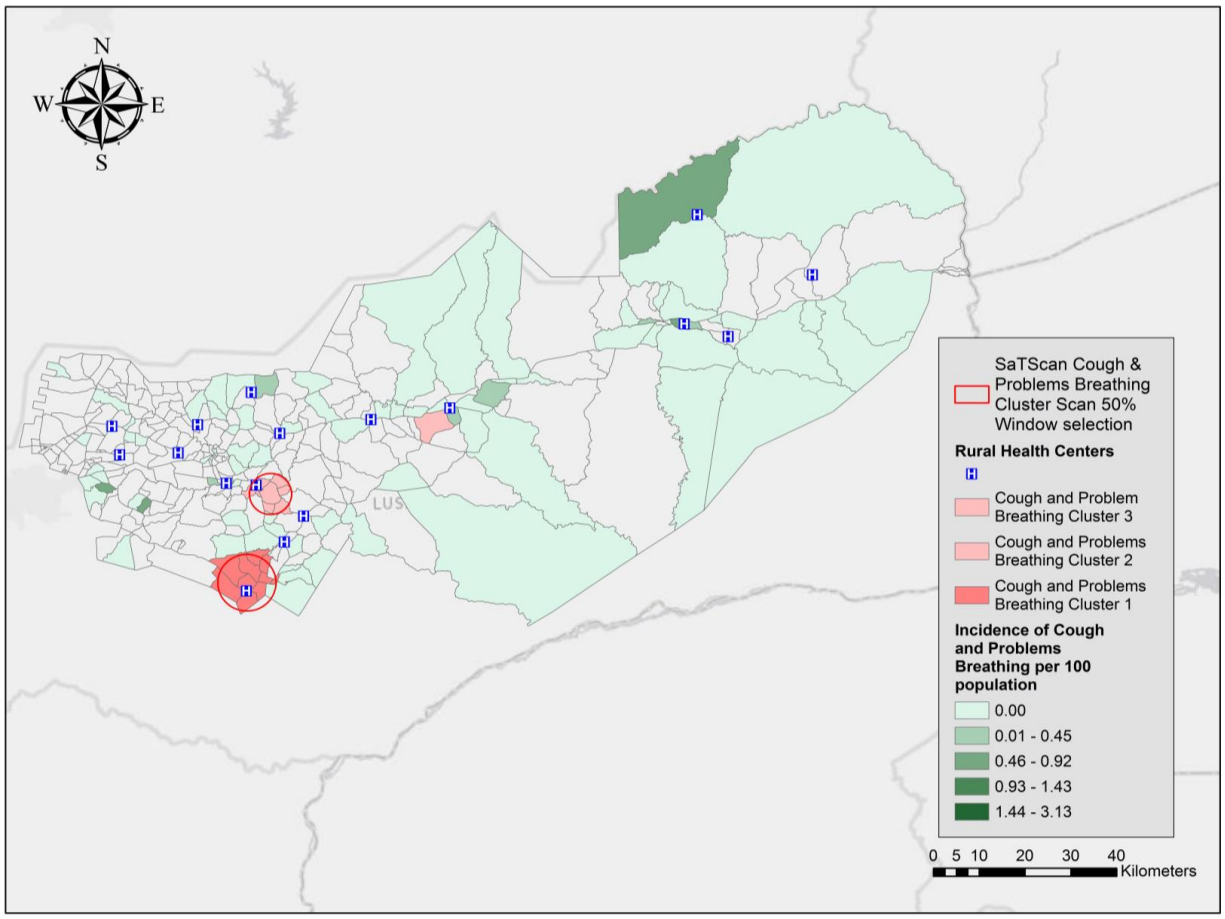


Figure 4.11 Pneumonia Incidence Cluster Areas

4.2.4.2 Severe Pneumonia

The cluster analysis of severe pneumonia illness incidence identified five primary and secondary cluster areas (Figure 4.12). The scan identified a primary cluster consisting of 10 SEA areas, which were the same areas that were identified in the cluster analysis for pneumonia. Two additional secondary clusters were also identified in this analysis. The Monte Carlo rank and RR were also the same as for pneumonia, indicating that the risk of severe pneumonia incidence is the same for pneumonia outcomes regardless of severity.

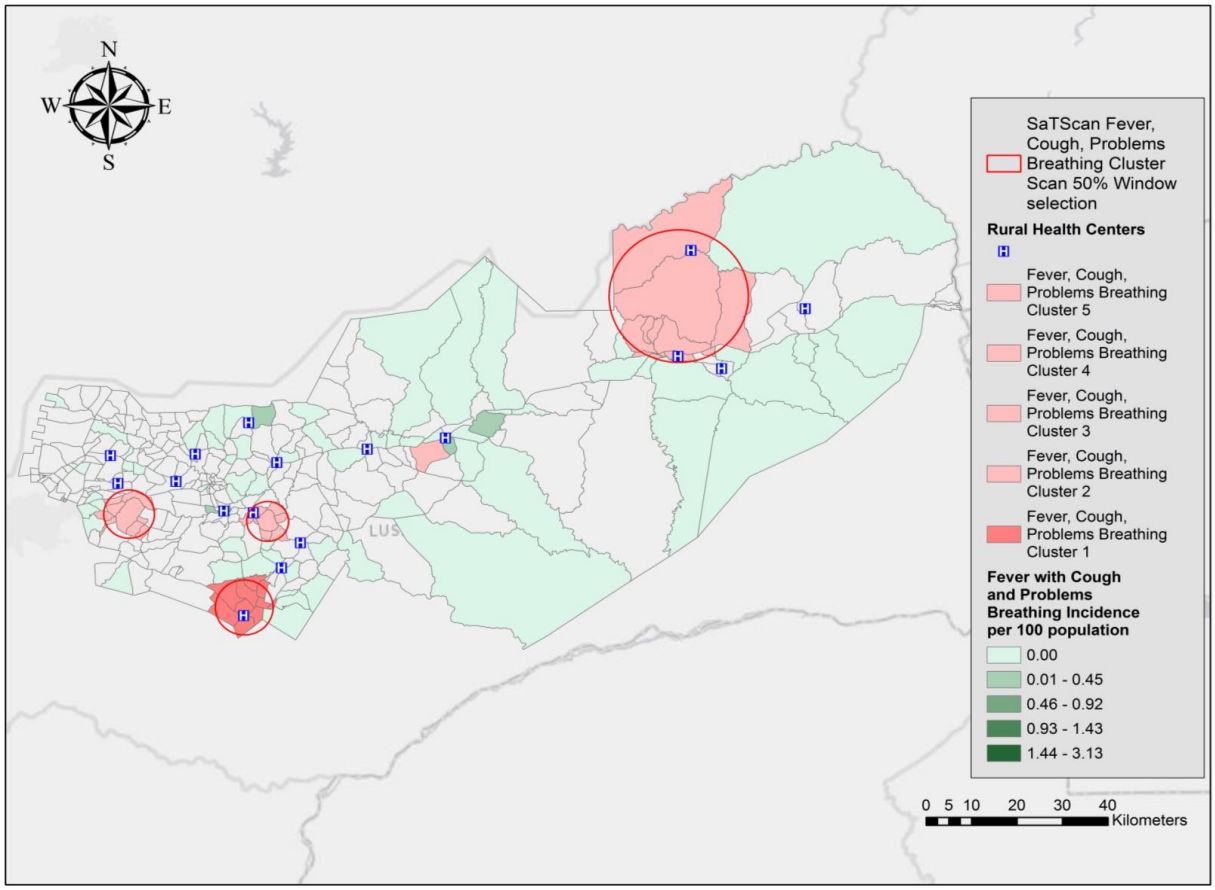


Figure 4.12 SaTScan v9.3 Severe Pneumonia Incidence Cluster Areas

4.3 Limitations

A number of limitations may affect the reliability of these results. Any final conclusions made should be made with caution and reservation. One of the primary limitations that should be noted is the quality of the dataset used to evaluate cluster areas. Geocoding methods were applied only to areas where village locations had been recorded by CHWs. If a CHW entered village names for a quarter of all patients in the area, this would lead to inaccurate assumptions about the incidence rate and RR calculations for that area. Secondly, validation of the matched village locations and ground truthing was not performed, with the possibility that villages may have been inaccurately matched to with the village coordinates or location area. It was assumed that village names that matched in both datasets were inherently the same village areas, given the same village nomenclature, indicated in the reference datasets. It is possible for villages to have the same name but be located in two entirely different locations. This would have an influence on the disease incidence in the SEAs. The locations that were identified serve as proxies for the true village location and should be considered to be the general location of a village and not the true location for all the villages in the dataset. As a result of using secondary data collection methods for the geolocations and not having complete census data for these locations, it was not possible to verify the accuracy for each village location coordinates with official government census. This is a primary limitation of secondary data analysis and should be noted for this study.

Third, the analyses were limited to purely spatial analysis for disease/illness outcomes without the consideration of time or age adjustments. A spatial-temporal analysis would provide insights into cluster variation of disease occurrence over time. A provision was made for these adjustments with the optional adjustment dataset file. The adjustment file only included adjustments for zero population areas, and even so, did not account for low population densities

and heterogeneity in risk was not accounted for. The Although CHW registries had a column for recording patients' age and dates of visits, too many inconsistencies and numerous missing data, were sufficient reasons to exclude these variables from the analyses. Future studies should address the importance of good record keeping and data completeness during CHW trainings.

Additional information on other covariates, such as gender and rainfall, would also have an influence on cluster outcomes and risk estimates. The outputs generated with very high RR should be further explored and tested using adjustments for known confounding and risk factors to adjust the relative risk rates to more accurate estimates for reliable inference and decision-making.

Another limitation was the issue of measurement error and misclassification bias. The CHWs data were not validated for accuracy or cross-referenced with follow-up visits, which could lead to measurement errors for cases of disease and illness. Also patients who had multiple visits recorded multiple times could lead to over reporting of cases. It was difficult to know how many times a patient was recorded with the same symptoms over multiple visits. However, because many of the illnesses that were observed are endemic and chronic, each occurrence of cases was treated independently regardless of the number of visits from the patient.

Lastly, it should also be noted that low and zero cases did not have an effect on the statistical stability of the risk estimates. SaTScan v9.3 analyses do not depend on the geographical resolution of the entered data, but rather on the population size of the circles constructed by SaTScan v9.3. It is a primary reason SaTScan v9.3 is used for health surveillance data, because it avoids arbitrary geographical aggregation of the data, allowing the scan statistic to consider varying sized aggregations through a moving window (Kulldorff 2014). High rates could then be attributed to confounding factors or covariates that would require further exploration.

CHAPTER FIVE: DISCUSSION

It was hypothesized that CHW health registries were an under utilized source of community health information, and that if health information were managed using mHealth devices quipped with GPS capabilities, then community health data could be analyzed for local area spatial variations using a GIS. It is thought that this would allow for more efficient health systems strengthening and would be advantageous to policy-makers, stakeholders, and communities. The remainder of this thesis will respond to the objective questions. Secondly an in-depth discussion on the findings of the study, which suggests the hypotheses to be valid, will be provided. The discussion on the primary objectives will be presented in the order that they were addressed in the research.

Can existing CHW health registries be used to create disaggregated sub-district level, low-resolution, geo-datasets suitable for geospatial health data analyses?

Yes, existing CHW health registries can be used to create sub-district level geo-datasets for Chongwe, Zambia. A number of methods for how to manually create a geodataset were explored to determine the best approach. A method that was explored required entering the names of the villages into Google Earth to identify the geolocations to determine the village locations manually. However, this approach yielded very little to no results. This was due to the occurrence of multiple spellings of village names in the CHW registry records, sometimes resulting in none of the spellings being truly accurate. So many village locations could not be located using Google Earth. Secondly, names of villages recorded may not have been the official names of the villages, often times the name of the headman of chief of a village was used as the village name and not the name listed for the official census, again resulting in the inability to find village location on Google Earth. Lastly, names of villages that were identified by Google Earth were locations for urban areas that had identical nomenclature. A village named “Kafue” was

identified as the city or region of Kafue in Google Earth and not the village itself. Another method that was explored for a manually generated geodataset relied on obtaining village coordinate locations from the Census Statistics Office or the Office of the Surveyor General to gain access to geodatasets containing coordinate information of village locations. This was done and it required an official written request from the Ministry of Health to the various directors of each agency. A request was made and after a month access was granted, however what was accessible was limited to what was readily available and processed by the agencies. Additionally, access to those sub-district level files was not available for public access and use. Unfortunately, there were costs associated with data acquisition as well. These costs were very expensive and in the end the decision was made to attempt other means. Fortunately, a few district level administrative area shapefiles obtained from CSO had village names and locations in the attribute tables. These attribute tables were used as a codebook for matching the village locations to the CHW registries.

Lastly, a third approach that was explored to identify village geo-locations was the use of a Gazetteer that could be downloaded freely from the Internet. This method, similar to a Google search was unsuccessful due to the same limitations in the variations of village names and the same name being used in multiple places. Also noted, were the poor spatial temporality of the Gazetteer, which was very outdated and did not include any new village areas created within the last 10 years.

Despite these challenges, an alternative approach was performed involving the use of secondary data obtained from existing NGO geodatabases and geodatasets. These data sources were used as reference geodatasets to identify the coordinate locations for each village, by matching village names to the reference geodatasets.

Village names recorded in the health registries were used in the geocoding methods for identifying patient village locations. The use of secondary geodatasets for use as georeferencing codebooks was the ideal method used to determine coordinate locations for each village. Once the locations were finalized, the dataset was entered into ArcGIS to confirm that the locations made sense in relation to where each CHW worked and the corresponding patient records.

The quality of this dataset depended on the completeness of the CHW registries. Ideally, CHWs who consistently recorded all the patient information would yield a higher quality geodataset. However in this case, there were a number of CHWs who had inconsistencies in their data, either in spelling the patient's address, or in recording the patient village name at all. In total, of the 7,673 patient records written in the registries, 2,529 did not have a village name recorded for the patients, and of the 5,144 that had a village name recorded, 3,130 were successfully geocoded by matching village names with secondary geodatabases as codebooks.

The final dataset was merged with administrative boundary area shapefiles and was suitable for further geospatial exploration and analyses for health outcomes at the sub-district levels at both the Ward and SEA levels.

Can the use of GIS on low-resolution health geo-datasets provide adequate spatial-temporal insights into CHW needs and community health outcomes at a sub-district level?

Yes, the final geodataset created was a complete digitization of paper-based CHW health registries managed at the sub-district community level. It served as a model for the type of output that could be created if mHealth devices were enabled with GPS capabilities for routine CHW activities.

The health registry geodataset had census data was added and population densities, case counts, and incidence rates were calculated. A visual exploration of the data was then performed. Through this visual exploration, the geographic distribution of villages and various incidence

cases were observed. This visual evaluation process was informative and provided insight on the village locations and community health burdens and the conditions in which CHWs worked. Looking at the preliminary maps it was clear that most health facilities were located in densely populated areas. It could be inferred that CHWs in these areas may have a larger volume of incident cases and may potentially experience frequent shortages in supplies and resources. For the CHWs who worked in remote areas, it was evident that the nearest RHC could be a great distance away, which could be a potential barrier for health access and service delivery. Many of the villages were located in densely populated areas near Lusaka city in the Western part of the district and along the main roadway, travelling Eastward, away from the Lusaka city. There was an area in Figure 5.1 located in the far Eastern part of the district, that had a high population density, indicated by the darker blue shade, located far from a nearby RHC. This would be an optimal area for policy-makers to explore for to establish new health facilities in the district.

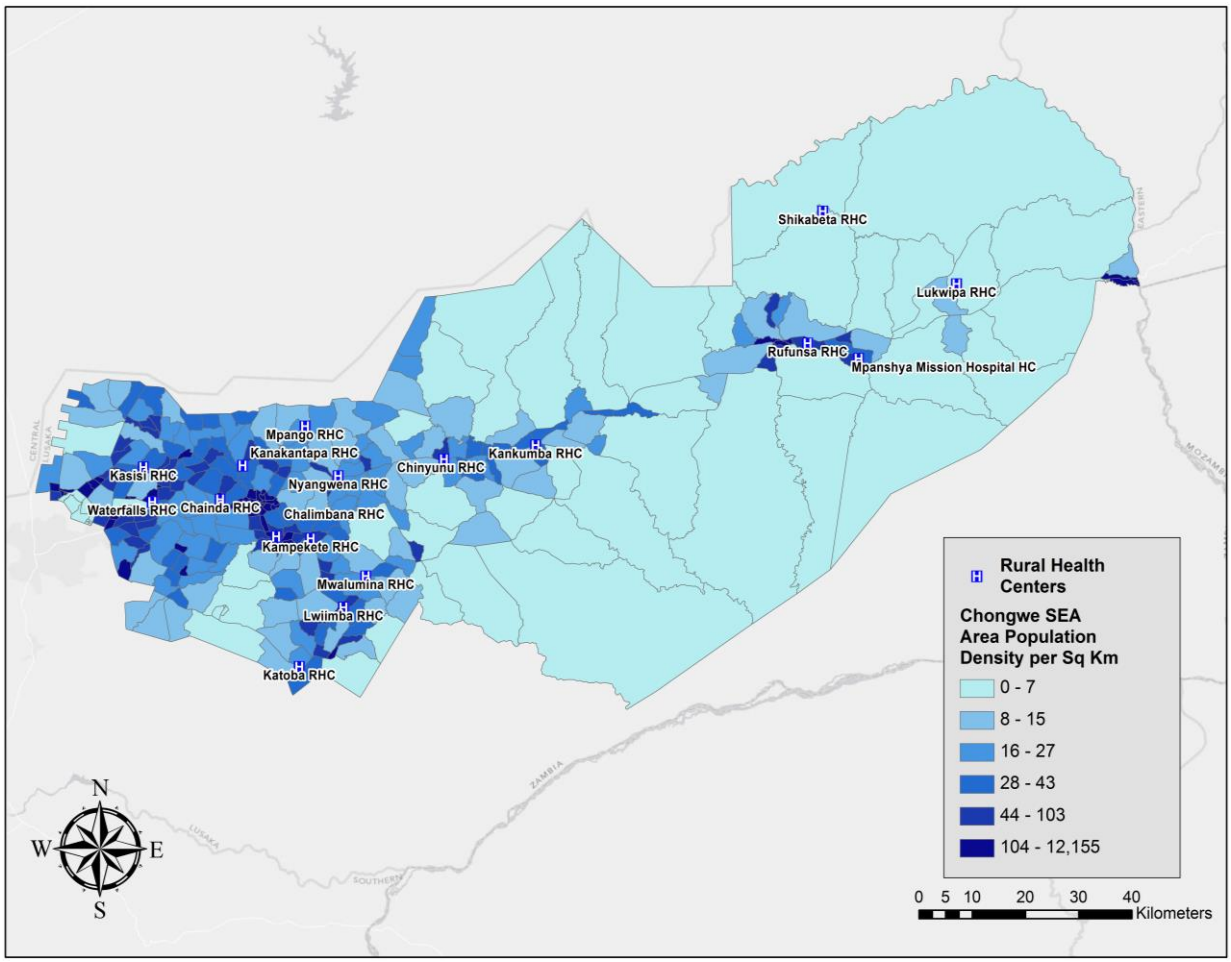


Figure 5.1 Population densities for Chongwe SEA areas and RHC locations.

Preliminary exploration of the disease/illness distribution maps also provided insights prior to analysis (Figure 5.2), on potential hot spot or cluster areas. A number of areas seemed to have a clustering in the number of illness cases. Some of the high numbers of cases identified on the map were in locations where a RHC was not nearby. Areas identified to have high case counts were noted as areas for additional exploration during the analyses. Additionally it was noted that there were areas with high fever counts but RDT usage was low, indicating potential areas of concern for misdiagnosis and mismanagement of malaria illness cases.

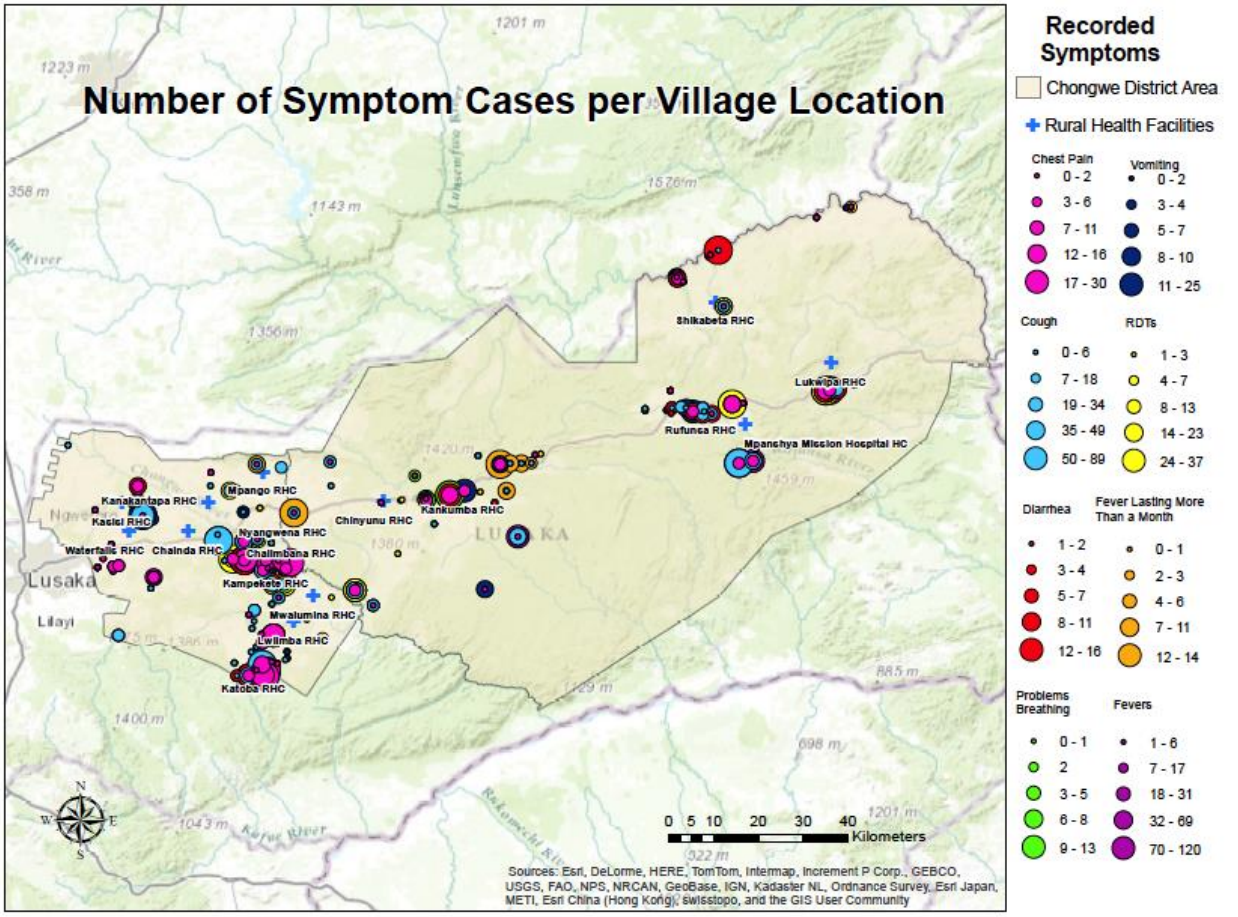


Figure 5.2 Exploratory map of disease/illness counts for each village location in relation to RHC

After visual exploration of the proportional density maps was complete, the next step was to determine if areas with high case counts were significant illness hot spots within the district. SaTScan v9.3 was used to conduct a cluster analysis. ArcGIS 10.2 had been suitable for visual explorations but was less ideal for performing a cluster analysis. The reason being, the geocoding methods used and the type of records available in the CHW health registries. Ideally if the cases had been individual case counts with unique geolocations, and not an aggregation into one village location point, ArcGIS could be used to perform a cluster analysis using Morans I for global district level cluster detection or Getis Ord Gi* for local sub-district cluster detection. However, the dataset consisted of binary data (yes and no) records for symptom case

management. This data type was not suitable for spatial analysis using the available ArcGIS tools and methods for hot spot and cluster analysis. There were a number of assumptions needed to run ArcGIS hot spot and cluster analyses that the GCC dataset failed to meet.

An alternative option for cluster detection analysis was the open source software SaTScan v9.3. The software uses population size and a moving scanning circular window of varying size to determine hot spots or clusters areas. A location that has a low number of cases is not influenced by the size of the underlying geographical area. An added benefit of the software is the adjustment file, which accounted for low or missing populations along with any covariates, and known risk ratios. The minimum number of spatial locations was also less restrictive than ArcGIS, the minimum number of locations needed to run a purely spatial scan was two locations, compared to 30 features for ArcGIS. Although for two locations a chi-square statistic could be used instead, but ideally for SaTScan v9.3, the more geographical locations the better.

To run the spatial analyses, SaTScan v9.3 required at minimum a case file for each disease/illness, a population file consisting of total population numbers for each geographic location, and a coordinate file with the latitude and longitude for each geographic location. Missing data could be adjusted for, including areas with zero population estimates, using an adjustment file.

For areas that were not included in the study area, resulting in no case counts, SaTScan v9.3 interpreted these areas as zero case counts, and these areas were included in the cluster analyses. This inclusion would most likely influence the RR calculated within the circular windows. Because the number of cases and the RR for those locations was unknown, any conclusions made from the analyses should be made with caution. High risks could be a result of low or missing case data or confounding effects not adjusted for. Additionally, the inconsistencies in recording village names by CHWs, also increases the level of caution to any

inferential conclusions about the risks. This is another limitation of this study, which can be accounted for if more CHW registries from more locations were included in the geocoding process. Inclusion of more CHW registries from more geographic locations would allow for a clearer assessment of the disease/illness distribution and risk estimates at the sub-district level.

Despite these limitations, areas were identified as clusters. The incidence and cluster analysis maps were shown to a panel of district health experts from the district, including the Provincial Medical Officer and the District Medical Officers. The maps were internally validated for reliability to confirm results matched what would be expected for identified cluster areas. The district medical officer and the principal investigators provided key informant information on the plausibility of cluster locations occurring in certain areas. Feedback provided by key informants revealed that pneumonia and fever clusters occurring in high incidence areas, were located in areas where CHW training for fever integrated case management had been inadequate, potentially leading to misdiagnosis or mismanagement of fevers in those areas. RDT usage was lowest in rural areas where malaria with fever was highest and RHC were few in the Eastern part of the district. It was noted that the Eastern part of the district was also an area that had been provided the fewest interventions to combat malaria and febrile illness. In the cluster areas for severe diarrhea, primary cluster areas occurred around Kampekete health center, a newly formed RHC in the district. This may have been an area in need of a health facility and the results support the need for one in the area. The development of a new health clinic may reduce severe diarrhea incidence in the area over time and this map could serve as a base line for future assessments. For malaria with fever incidence and clustering, there were a few locations identified as primary cluster sites in densely populated rural areas in the Western part of the district. Health experts cited the locations of two nearby rivers in the area, providing an opportune vector area for mosquitoes and increased risk for malaria. Further exploration in

environmental controls in the area could lead to decreases in incidence and elimination of the cluster areas. Overall, the identified cluster areas detected by SaTScan v9.3 coincided with expectations provided by health experts and key informants interviewed who were familiar with the area.

Are CHW community health registries an under-used source of information for insights into sub-district level standards of care and health outcomes?

Based on the findings of this thesis, CHW health registries are an underused data source capable of providing beneficial health information at the sub-district and community levels. Coincidently, many mHealth programs do not include routine CHW case management registry indicators into mHealth device programming. Instead most mHealth programs are designed to address or explore a specific case management area like malaria, HIV, and maternal child health indicators, etc alone. The beneficial insights that can be gained by evaluating illnesses at the sub-district level, using mHealth methods, would be advantageous towards understanding the epidemiology of diseases and illness within specific areas. Additionally, mHealth programs that are coupled with GIS would strengthen surveillance and monitoring of diseases/illnesses, CHW performance and impacts, and the efficacy of implemented health programs within communities.

Current data management processes attenuate the underlying epidemiology of many disease/illnesses occurring within a district in Zambia. Data collected by CHWs is aggregated to the nearest RHC, then from the RHCs the data is aggregated further to the district health centers, once at the district health centers the data is sent to the provincial health centers who then report data for the district to the Ministry of Health located at the central level. This process is time consuming and potentially masks any hot spots or cluster areas occurring in a sub-district area. For Chongwe, 334 SEA areas where located within the district and 102 were analyzed for in the

analyses. Performing an analysis on approximately one-third of the SEA areas identified cluster areas and provided other useful and beneficial information, suggesting that data at the sub-district level should be incorporated into health systems analyses and surveillance for more timely and accurate assessments.

Community health worker health registries are a primary source for community data and would be ideal for sub-district community level surveillance. The emergence of mHealth and existing GIS technology create an ideal opportunity to begin using non-aggregated CHW generated health data for evidence-based decision making addressing monitoring and treatment of disease outcomes. Doing so would increase efficacy and timeliness of outbreak response and other interventions. Community health worker registries are indeed an overlooked data source and an under utilized asset.

Significance for Other Low-resource Settings and Regions

For other low-resource settings similar to Zambia, this project illustrates not only the importance of geospatial thinking in disease surveillance and disease management, but also that barriers associated with GIS technology costs can be mitigated. The methods used for the retrospective analysis of the CHW patient registries can be applied to any registry or database that had a geographic place or name that can be identified on a map. Other areas in which these methods can be applied are Road Traffic Accident studies. Accident locations recorded in paper-based police registries (ex. Kafue Rd 2 miles from Chongwe village) could be mapped using similar methods and techniques presented in this paper. By working in collaboration with local agencies that routinely collect geospatial data, manual geocoding methods can be achieved, as was done with the Central Statistic Office or PATH. Additionally it provides an opportunity for more remote settings to be mapped and databases generated. The geocoding process may take time, but the benefits outweigh the cost in time or in not doing so at all.

Community-based participatory mapping is another way that coordinate locations could be gathered at a low cost. This low-tech approach involves creating large paper-based maps with latitude and longitude degree grid markings of a community area. Community members are then asked to place stickers of varying shapes, sizes, and colors on top of the map at locations that they know, this can include households, schools, landmarks, etc. This coordinate data can then be entered into a GIS or Google Map for visualization and further analysis. In this project, barriers for geocoding were overcome through partnerships with NGOs and other government agencies, similar approaches can be done in other settings.

Secondly, the use of an open-sourced software also illustrates way in which GIS cost barriers can be overcome. There have been a number of open-sourced GIS programs that are available for data analysis and mapping. One that is most frequently used is QGIS, which is completely free and is capable of performing the same data analysis of commonly used proprietary software. There are a number of free tutorials and online resources to learn how to use the software.

Lastly, the main take-away from this research is that geospatial thinking and implementation can be achieved without the need of expensive technology or software. In many developing regions, paper-based methods are still the primary form for data management. Even with these practices and norms, geospatial methods can be applied as long as a geography component exists in the data. The availability of open-sourced software also opens up more opportunities to overcome financial barriers. Geospatial thinking enhances the depth of knowledge in which decisions are made and if time or resources are limited, this is a good way to prioritize decisions on policies and actions.

CHAPTER SIX: CONCLUSIONS

In conclusion, the findings of this study support the concept of applying GIS to CHW registries. The model dataset yielded beneficial health information and provided geospatial insight into the distribution of disease/illness at the community level. It also provided support for mHealth initiatives to incorporate both routine CHW indicators into mHealth device programming for use with a GIS. By coupling mHealth initiatives with GIS, it provides opportunities to explore disease incidence over space and time, further encompassing the epidemiological tenets of person, place, and time. Thus resulting in more efficient and accurate epidemiological assessments within the district.

6. 1 Policy Implications

The use of mHealth and GIS on CHW health registries creates an opportunity for more informed decision-making, through evidence-based support for community interventions and policy development. Analysis of sub-district level disease/illness distributions would have immediate impact on health policies at the community level. Using GIS to identify cluster areas would provide evidence for more targeted interventions and strategic roll out of implemented programs. Additionally, stakeholders and community developers could make better-informed decisions on where new RHC would be most beneficial using incidence and density maps.

Routine surveillance and maps of sub-district areas would also inform stakeholders and policy makers on programmatic impact and efficacy. In the case of severe diarrhea, the cases were clustered around a new health facility. With routine community level surveillance and mapping, changes in the cluster area could be monitored after additional CHW trainings and health interventions are employed. Changes in size and presence of the cluster altogether would inform if the interventions were efficacious. In areas where malaria and fever clustered, policy makers could make informed decisions on where limited RDT resources should be allocated.

Also, monitoring and impact evaluations could occur after RDT resources are allocated and malaria interventions are implemented in these areas.

The immediate impact these findings have on policies in Zambia is that they provide some evidence to support the claim that the current system for reporting health data at district and provincial level is insufficient. By aggregating community level data, many local variations in disease outcomes, especially in rural areas, are inadequately reported reducing the surveillance potential for rapid response to any disease/illness outbreaks. The primary policy recommendation and next steps, based on these findings, would be to integrate GIS into existing health systems for more accurate sub-district level data acquisition and real-time surveillance.

The recent introduction of the District Health Information System version 2 ($DHIS_2$) provides an ideal opportunity to implement this recommendation immediately at a national level. The $DHIS_2$ is a health information system used to manage data collected at the district levels from all health facilities in a district. This second generation system has mapping and geospatial data management capabilities integrated into the framework. However current implementation of the system has yet to install or activate this feature in many of the districts, based on key informant interviews. The system is able to visualize geospatial data down to the health facility level for all health facilities in a district for real-time disease surveillance. The ability to acquire and visualize data at a sub-district level opens up opportunities to merge mHealth initiatives using CHW community level data directly into the $DHIS_2$ system. This would optimize the performance of the system and allow for user-generated data to be monitored immediately for surveillance purposes, especially in rural areas and areas hard to travel to in the districts. Additionally, data could be downloaded for further geospatial analysis by facility level, or community level if merged with mHealth data, providing evidence for more informed decision-making and policy development relevant to individual communities. Until the $DHIS_2$ mapping

capabilities are fully activated to include sub-district data from individual health facilities, the system will continue to be under utilized and disease surveillance inadequate for many communities in Zambia. The findings of this research illustrate the advantages of looking at diseases and illness through a geospatial lens and the benefits of geospatial analysis for decision-making and policy development.

6.2 Future Research

Future research should evaluate cluster areas for variations in age, gender, and over time. By stratifying analyses by age or gender, additional sub-district variations could be identified for vulnerable populations such as children under-five and pregnant women. Due to inconsistencies and missing data, a cluster analyses by age and gender could not be done with the existing CHW registry dataset. Space-time cluster analyses could not be performed as a result of incomplete and inconsistent data entry. This is often the case with paper-based data collection methods and secondary data analyses.

Future mHealth research should also explore coupling with GPS and integrating CHW registry indicators into mHealth device programming. By incorporating the use of mHealth technology for CHW community health data management, internal quality assurance features could be implemented to prevent incomplete or inconsistent data collection by CHWs. This would also provide insights into individual CHW impact and would identify areas for additional CHW trainings. It would also result in higher quality datasets for more types of analyses, such as a geographic weighted regression, to identify spatial relationships related to elevation, rainfall, temperature, and landuse, etc.

A cost benefit analysis should also be explored to identify the benefits and costs of broad implementation of mHealth programs. In addition to this, explorations into various types of mHealth devices should also be explored. As smart phone technology becomes more available

globally and cloud based data management becomes more prevalent, insights into performance and outcomes related to the types of devices used would be interesting to know and would help guide the types of devices that are best for mHealth utilization in resource poor settings.

Finally, validation of the findings of this initial assessment should be done using primary data collection methods of village locations and additional CHW registry datasets throughout the district.

REFERENCES

Abrams, A.M., M. Kleinman, and M. Kulldorff. 2014. Gumbel based p-value approximations for spatial scan statistics. *International Journal of Health Geographics* 9: 61. http://www.ij-healthgeographics.com/content/9/1/61 (last accessed 29 August 2014).

Acestor, N., R. Cooksey, P. N. Newton, D. Ménard, P. J. Guerin, J. Nakagawa, E. Christophel, I. J. González, and D. Bell. 2012. Mapping the aetiology of non-malarial febrile illness in Southeast Asia through a systematic review--terra incognita impairing treatment policies. *PLoS ONE* 7 (9): e44269.

Amexo, M., R. Tolhurst, G. Barnish, and I. Bates. 2004. Malaria misdiagnosis: effects on the poor and vulnerable. *Lancet* 364 (9448): 1896–8.

Bell, B.S., R.E. Hoskins, L.W. Pickle, and D. Wartenberg. 2006. Current practices in spatial analysis of cancer data: mapping health statistics to inform policymakers and the public. *International Journal of Health Geographics* 5: 49. http://www.ij-healthgeographics.com/content/5/1/49 (last accessed 29 August 2014).

Black, R.E., S. Cousens, H.L. Johnson, J.E. Lawn, I. Rudan, D.G. Bassani, P. Jha, H. Campbell, C.F. Walker, R. Cibulskis, T. Eisele, L. Liu, and C. Mathers. 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 375 (9730): 1969–87.

Blaya, J., H. Fraser, and B. Holt. 2010. E-health technologies show promise in developing countries. *Health Affairs* 2 (2): 244–251.

Braun, R., C. Catalani, J. Wimbush, and D. Israelski. 2013. Community health workers and mobile technology: a systematic review of the literature. *PLoS ONE* 8 (6): e65772.

Ceccato, P., S.J. Connor, I. Jeanne, and M.C. Thomson. 2005. Application of Geographical Information Systems and Remote Sensing technologies for assessing and monitoring malaria risk. *Parassitologia* 47 (1): 81–96.

Central Statistical Office. 2012a. *2010 Census of Population and Housing National Descriptive Tables*.

———. 2012b. *2010 Census of Population and Housing National Descriptive Tables Volume 11*.

———. 2012c. *2010 Census of Population Summary Report*.

Chanda, E., and V. Mukonka. 2012. Using a geographical-information-system-based decision support to enhance malaria vector control in Zambia. *Journal of Tropical Medicine* 2012: 363520

Chanda, P., B. Hamainza, H.B. Moonga, V. Chalwe, and F. Pagnoni. 2011. Community case management of malaria using ACT and RDT in two districts in Zambia: achieving high adherence to test results using community health workers. *Malaria Journal* 10: 158. http://www.malariajournal.com/content/pdf/1475-2875-10-158.pdf (last accessed 29 August 2014).

Chanda, P., B. Hamainza, and S. Mulenga. 2009. Malaria control and implications for the management of fever in under-five children at a peripheral health facility: a case study of Chongwe rural health centre in. *Malaria Journal* 8: 49. http://www.malariajournal.com/content/8/1/49 (last accessed 1 September 2014).

Chipwaza, B., J.P. Mugasa, I. Mayumana, M. Amuri, C. Makungu, and P.S. Gwakisa. 2014. Community knowledge and attitudes and health workers' practices regarding non-malaria febrile illnesses in eastern Tanzania. *PLoS Neglected Tropical Diseases* 8 (5): e2896. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4031176 (last accessed 30 August 2014).

Clemmer, G. 2010. *The GIS 20 essential skills* 1rst ed. Redlands, California: ESRI Press.

Cromley, E.K. 2003. GIS and Disease. *Annual Review of Public Health* 24: 7–24.

Cromley, E.K., and S.L. McLafferty. 2012. *GIS and public health* 2nd ed. New York, New York, USA: The Guilford Press.

d'Acremont V., A. Malila, N. Swai, R. Tillya, J. Kahama-Maro, C. Lengeler, and B. Genton. 2010. Withholding antimalarials in febrile children who have a negative result for a rapid diagnostic test. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 51 (5): 506–11.

DeRenzi, B., B. Birnbaum, L. Findlater, J. Mangilima, J. Payne, T. Parikh, G. Borriello, and N. Lesh. 2012. Improving community health worker performance through automated SMS. *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12*: 25. http://dl.acm.org/citation.cfm?doid=2160673.2160677.

Derua, Y.A., D.R. Ishengoma, R.T. Rwegoshora, F. Tenu, J.J. Massaga, L.E. Mboera, and S.M. Magesa. 2011. Users' and health service providers' perception on quality of laboratory malaria diagnosis in Tanzania. *Malaria Journal* 10: 78. http://www.malariajournal.com/content/10/1/78 (last accessed 30 August 2014).

English, M., H. Reyburn, C. Goodman, and R.W. Snow. 2009. Abandoning presumptive antimalarial treatment for febrile children aged less than five years--a case of running before we can walk? *PLoS medicine* 6 (1): e1000015. http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.1000015 (last accessed 25 November 2013).

Fisher, R.P., and B.A. Myers. 2011. Free and simple GIS as appropriate for health mapping in a low resource setting: a case study in eastern Indonesia. *International Journal of Health Geographics* 10: 15. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3051879/pdf/1476-072X-10-15.pdf (last accessed 11 November 2012).

Freifeld, C.C., R. Chunara, S.R. Mekaru, E.H. Chan, T. Kass-Hout, A. Ayala Iacucci, and J.S. Brownstein. 2010. Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS medicine* 7 (12): e1000376. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2998443 (last accessed 21 August 2014).

Gosling, R.D., C.J. Drakeley, A. Mwita, and D. Chandramohan. 2008. Presumptive treatment of fever cases as malaria: help or hindrance for malaria control? *Malaria Journal* 7: 132. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2488354 (last accessed 26 November 2013).

Hamer, D.H., M. Ndhlovu, D. Zurovac, M. Fox, K. Yeboah-Antwi, P. Chanda, N. Sipilinyambe, J.L. Simon, and R.W. Snow. 2007. Improved diagnostic testing and malaria treatment practices in Zambia. *JAMA : The Journal of the American Medical Association* 297 (20): 2227–31.

Hastings, I. 2011. How artemisinin-containing combination therapies slow the spread of antimalarial drug resistance. *Trends in Parasitology* 27 (2): 67–72.

Hume, J.C.C., G. Barnish, T. Mangal, L. Armázio, E. Streat, and I. Bates. 2008. Household cost of malaria overdiagnosis in rural Mozambique. *Malaria Journal* 7: 33. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2279141/pdf/1475-2875-7-33.pdf (last accessed 26 November 2013).

Kalyango, J.N., T. Alfven, S. Peterson, K. Mugenyi, C. Karamagi, and E. Rutebemberwa. 2013. Integrated community case management of malaria and pneumonia increases prompt and appropriate treatment for pneumonia symptoms in children under five years in Eastern Uganda. *Malaria Journal* 12: 340. http://www.malariajournal.com/content/12/1/340 (last accessed 13 October 2013).

Kamanga, A., P. Moono, G. Stresman, S. Mharakurwa, and C. Shiff. 2010. Rural health centres, communities and malaria case detection in Zambia using mobile telephones: a means to detect potential reservoirs of infection in unstable transmission conditions. *Malaria Journal* 9: 96. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2861692 (last accessed 16 October 2013).

Kandala, N.B., M.A. Magadi, and N.J. Madise. 2006. An investigation of district spatial variations of childhood diarrhoea and fever morbidity in Malawi. *Social Science & Medicine* 62 (5): 1138–52.

Kazembe, L.N., A.S. Muula, C.C. Appleton, and I. Kleinschmidt. 2007. Modelling the effect of malaria endemicity on spatial variations in childhood fever, diarrhoea and pneumonia in Malawi. *International Journal of Health Geographics* 6: 33. http://www.ij-healthgeographics.com/content/6/1/33 (last accessed 29 August 2014).

Kulldorff, and Martin. 2014. SaTScan User Guide Version: SaTScan v9.3 released March 20 2014. http://www.satscan.org/techdoc.html (last accessed 29 August 2014).

Manyando, C., E.M. Njunju, J. Chileshe, S. Siziya, and C. Shiff. 2014. Rapid diagnostic tests for malaria and health workers' adherence to test results at health facilities in Zambia. *Malaria Journal* 13: 166. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4026818 (last accessed 29 August 2014).

Msellem, M.I., A. Mårtensson, G. Rotllant, A. Bhattarai, J. Strömberg, E. Kahigwa, M. Garcia, M. Petzold, P. Olumese, A. Ali, and A. Björkman. 2009. Influence of rapid malaria diagnostic tests on treatment and health outcome in fever patients, Zanzibar: a crossover validation study. *PLoS medicine* 6 (4): e1000070.

Nhavoto, J.A., and A. Grönlund. 2014. Mobile technologies and geographic information systems to improve health care systems: a literature review. *JMIR mHealth and uHealth* 2 (2): e21. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4114429 (last accessed 24 August 2014).

O'Dempsey, T.J., T.F. McArdle, B.E. Laurence, A.C. Lamont, J.E. Todd, and B.M. Greenwood. 1993. Overlap in the clinical features of pneumonia and malaria in African children. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 87 (6): 662–5.

O'Sullivan, D., and D.J. Unwin. 2010. *Geographic information analysis* 2nd ed. Hoboken, New Jersey: John Wiley & Sons Inc.

Price, M. 2010. *Mastering ArcGIS* 4th ed. New York, New York, USA: McGraw-Hill.

Rafael, M., T. Taylor, A. Magill, Y. Lim, F. Girosi, and R. Allan. 2006. Reducing the burden of childhood malaria in Africa: the role of improved. *Nature* 444 (Suppl 1): 39–48.

Rao, V., D. Schellenberg, and A. Ghani. 2013. Impact of improving appropriate treatment for fever on malaria and non-malarial febrile illness management in under-5s: a decision-tree modelling approach. *PLoS ONE* 8 (7): e69654. http://dx.plos.org/10.1371/journal.pone.0069654.g005 (last accessed 1 September 2014).

Republic of Zambia Ministry of Health. 2011a. *Annual Health Statistical Bulletin*.

———. 2011b. *National Health Strategic Plan 2011-2015*.

———. 2012. *National Malaria Control Programme Strategic Plan for FY 2011-2015*.

Slater, H., and E. Michael. 2012. Predicting the current and future potential distributions of lymphatic filariasis in Africa using maximum entropy ecological niche modelling. *PLoS ONE* 7 (2): e32202. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3281123 (last accessed 22 August 2014).

Snow, R.W., K. Marsh, and D. le Sueur. 1996. The need for maps of transmission intensity to guide malaria control in Africa. *Parasitology Today* 12 (12): 455–457.

Tanser, F.C., and D. le Sueur. 2002. The application of geographical information systems to important public health problems in Africa. *International Journal of Health Geographics* 1: 4. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=149399 (last accessed 25 August 2014).

UNAIDS. Zambia Country Profile. http://www.unaids.org/en/regionscountries/countries/zambia/ (last accessed 5 August 2014).

United Nations Children's Fund. 2013. Malaria. *UNICEF Zambia Fact Sheets*. http://www.unicef.org/zambia/5109_8454.html (last accessed 15 November 2013).

United States of America Central Intelligence Agency. 2014. CIA World Factbook Africa: Zambia. *The World Factbook*. https://www.cia.gov/library/publications/the-world-factbook/geos/za.html (last accessed 22 April 2014).

White, L.J., P.N. Newton, R.J. Maude, W. Pan-ngum, J.R. Fried, M. Mayxay, R.R. Maude, and N.P.J. Day. 2012. Defining disease heterogeneity to guide the empirical treatment of febrile illness in resource poor settings. *PLoS ONE* 7 (9): e44545.

WHO. 2013. *World Malaria Report 2013*. http://www.who.int/malaria/publications/world_malaria_report_2013/en/ (last accessed 21 August 2014).

———. 2014. Zambia Country Profile. http://www.who.int/countries/zmb/en/ (last accessed 20 March 2014).

World Health Organization. 2013a. Children: Reducing Mortality Fact Sheet No. 178. *Media Centre* http://www.who.int/mediacentre/factsheets/fs178/en/ (last accessed 19 November 2013).

———. 2013b. Diarrheal Disease Fact Sheet No. 330. *Media Centre* http://www.who.int/mediacentre/factsheet/fs330/en/ (last accessed 19 November 2013).

Young, M., C. Wolfheim, D.R. Marsh, and D. Hammamy. 2012. World Health Organization/United Nations Children's Fund joint statement on integrated community case management: an equity-focused strategy to improve access to essential treatment services for children. *The American Journal of Tropical Medicine and Hygiene* 87 (5 Suppl): 6–10.

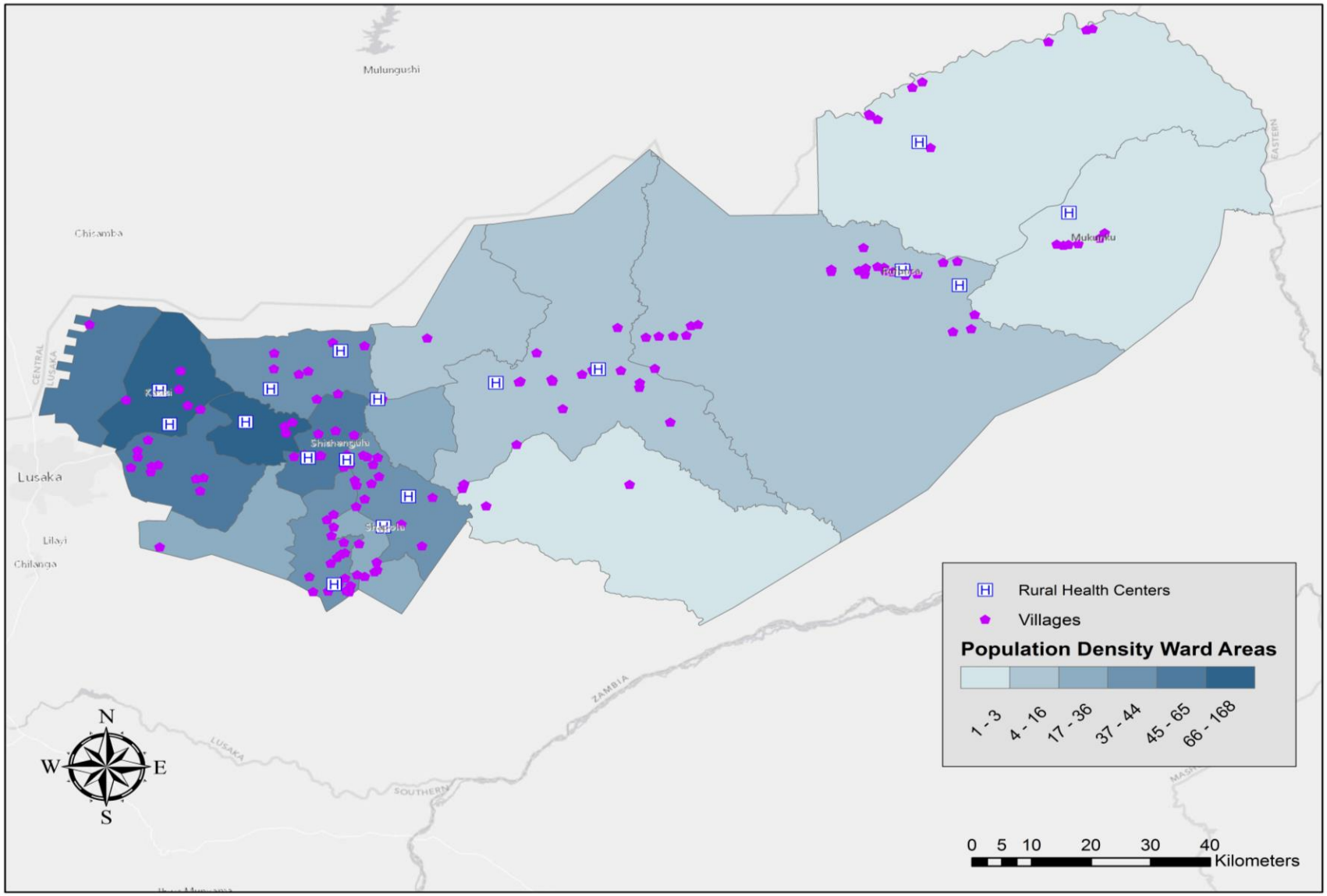
Zambia Tourism Board. Official Zambia Tourism Website. http://www.zambiatourism.com (last accessed 25 January 2014).

Zeiler, M. 1999. *Modeling our world* 1rst ed. Redlands, California: ESRI Press

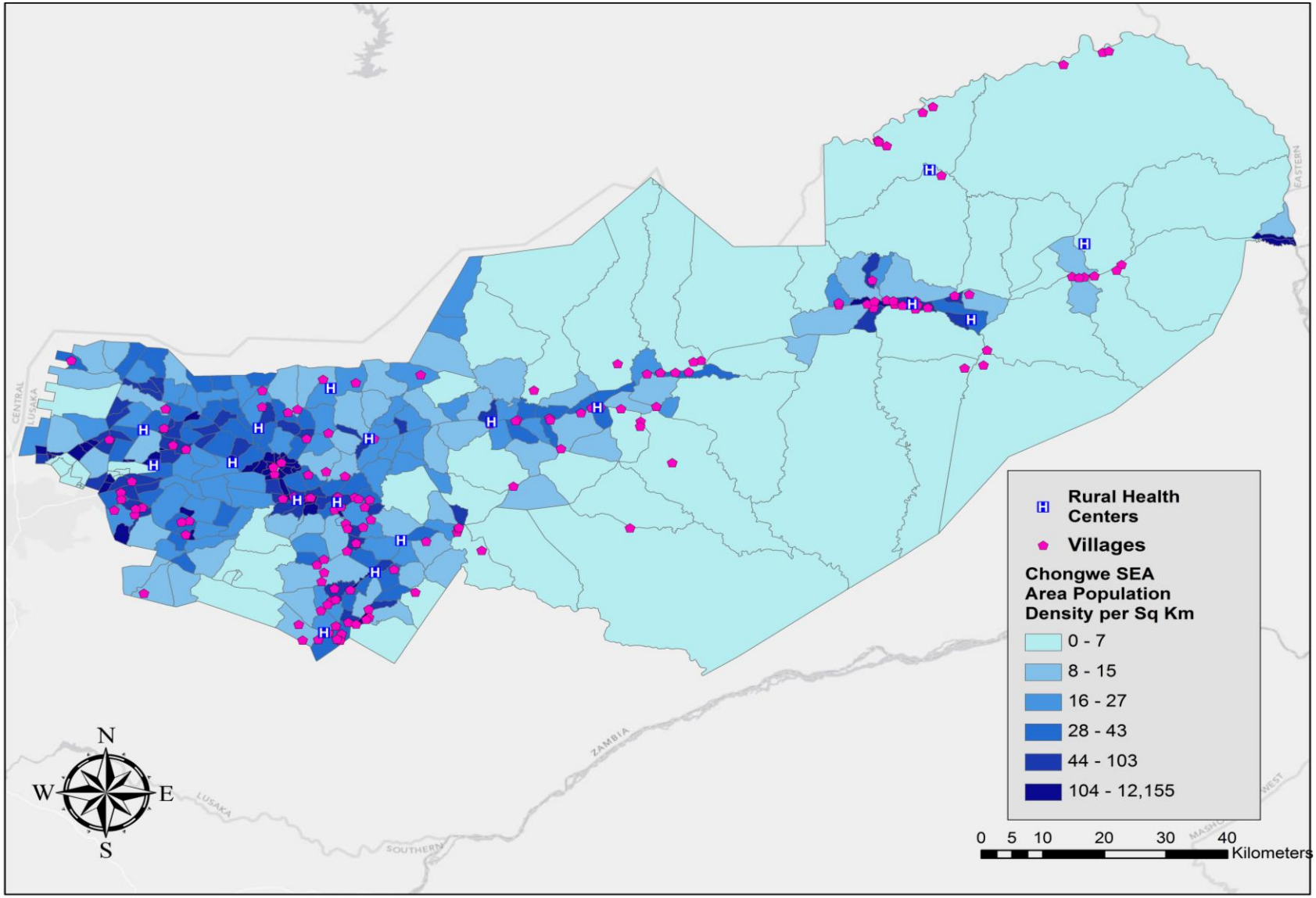
APPENDICES

Appendix A: Population Density Maps

Chongwe district population densities for Ward administrative areas with village and RHC locations

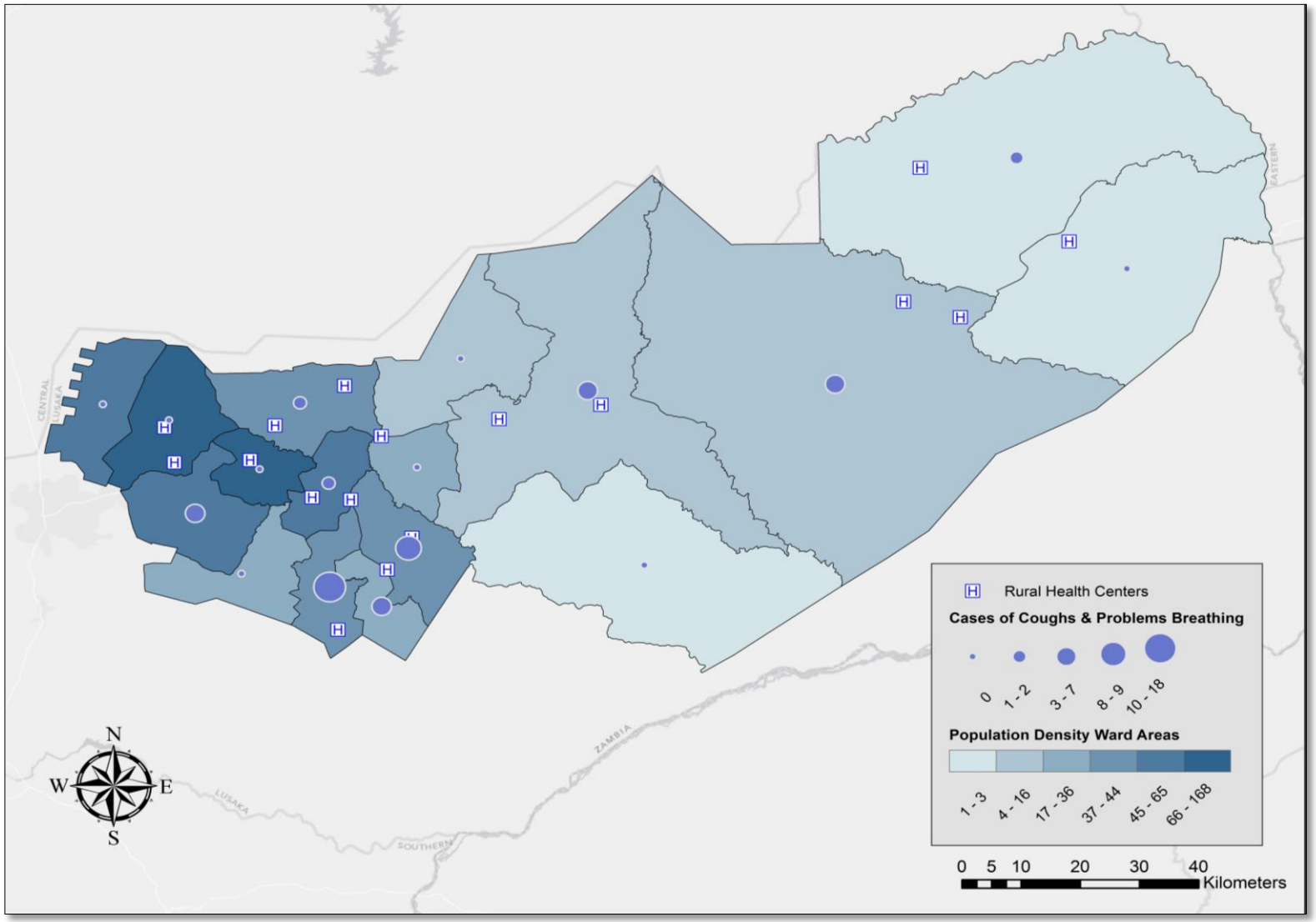


Chongwe population densities for SEA administrative boundary areas and RHC locations

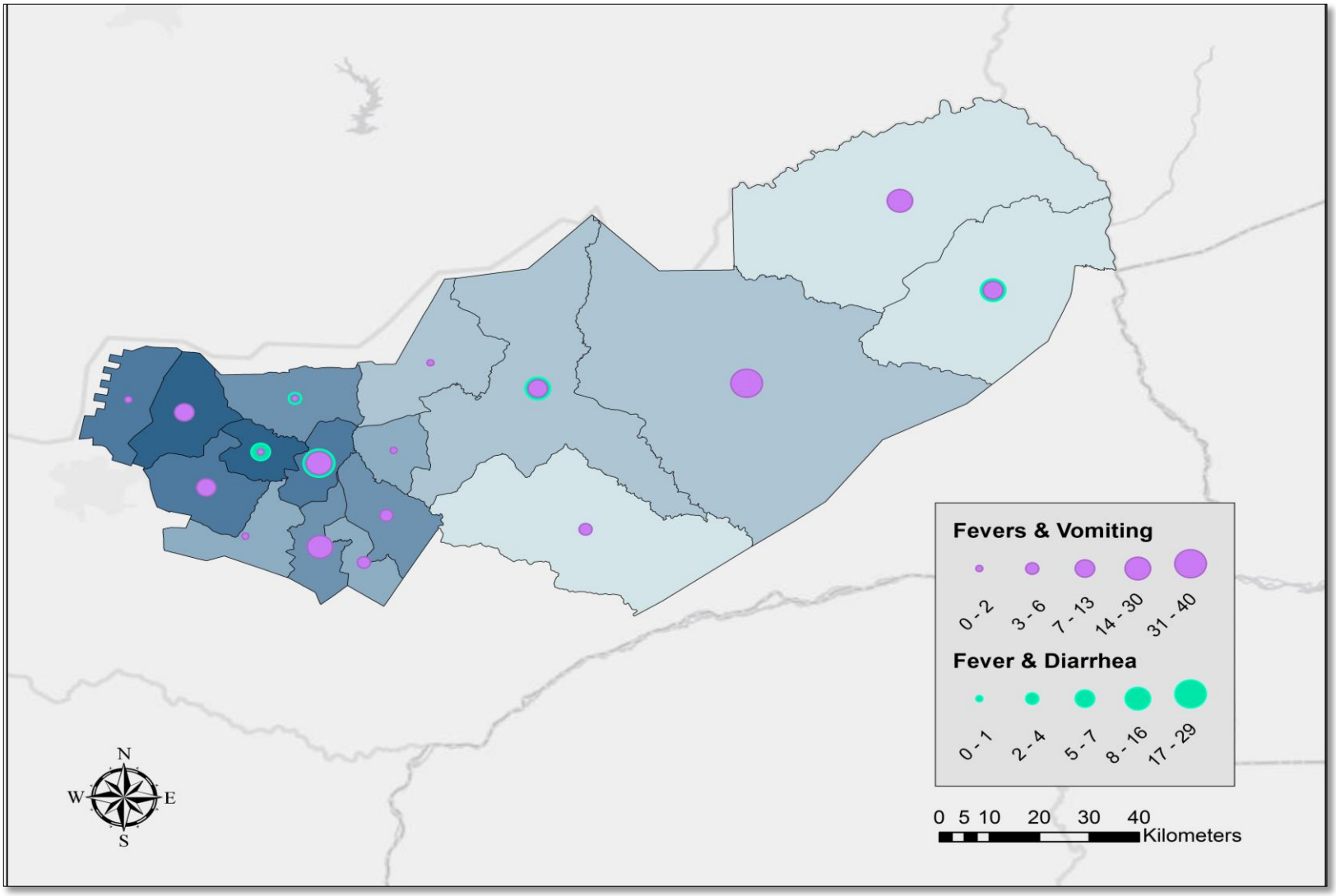


Appendix B: Disease/Illness Count Proportional Symbols Maps

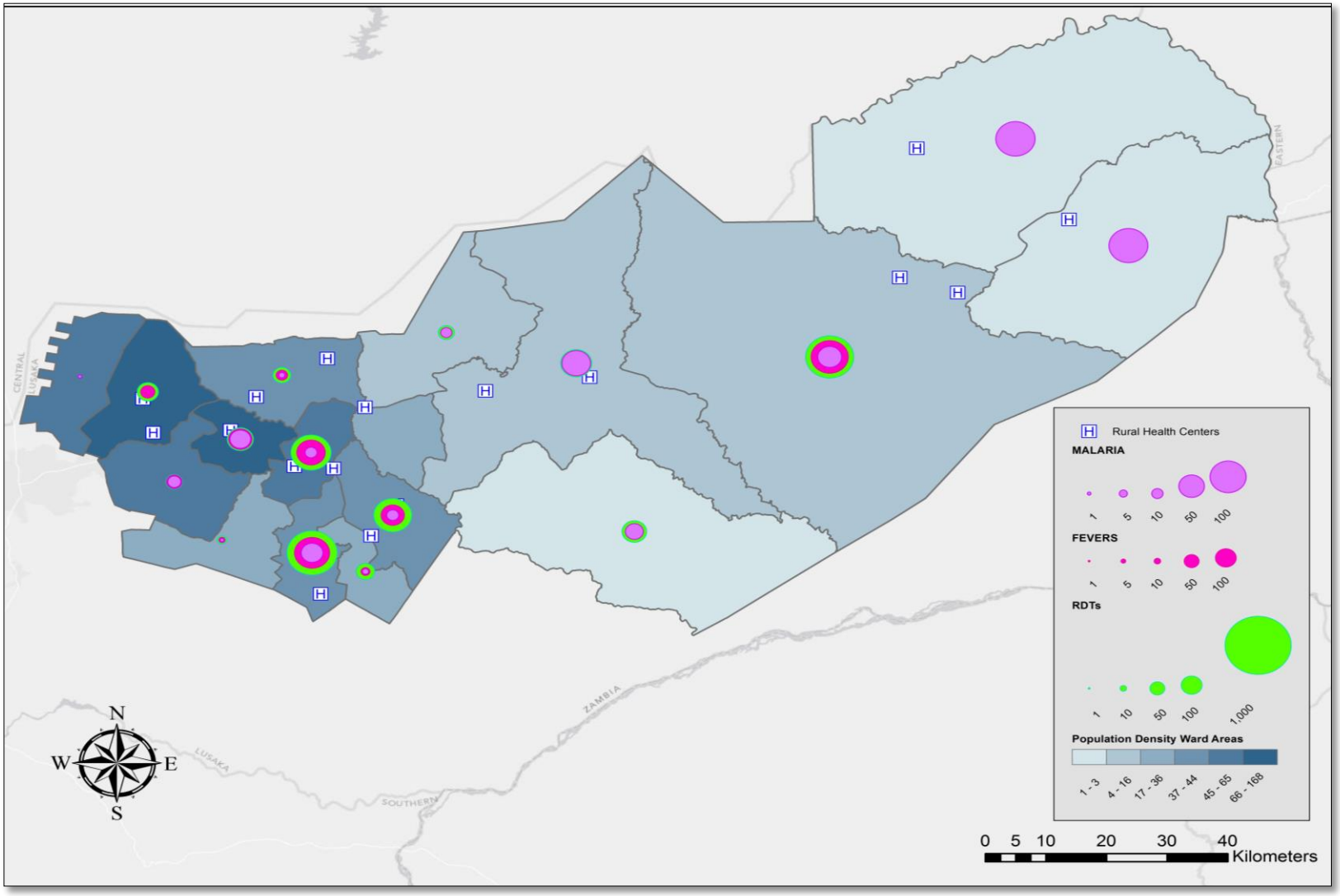
Cases of Coughs with Problems Breathing



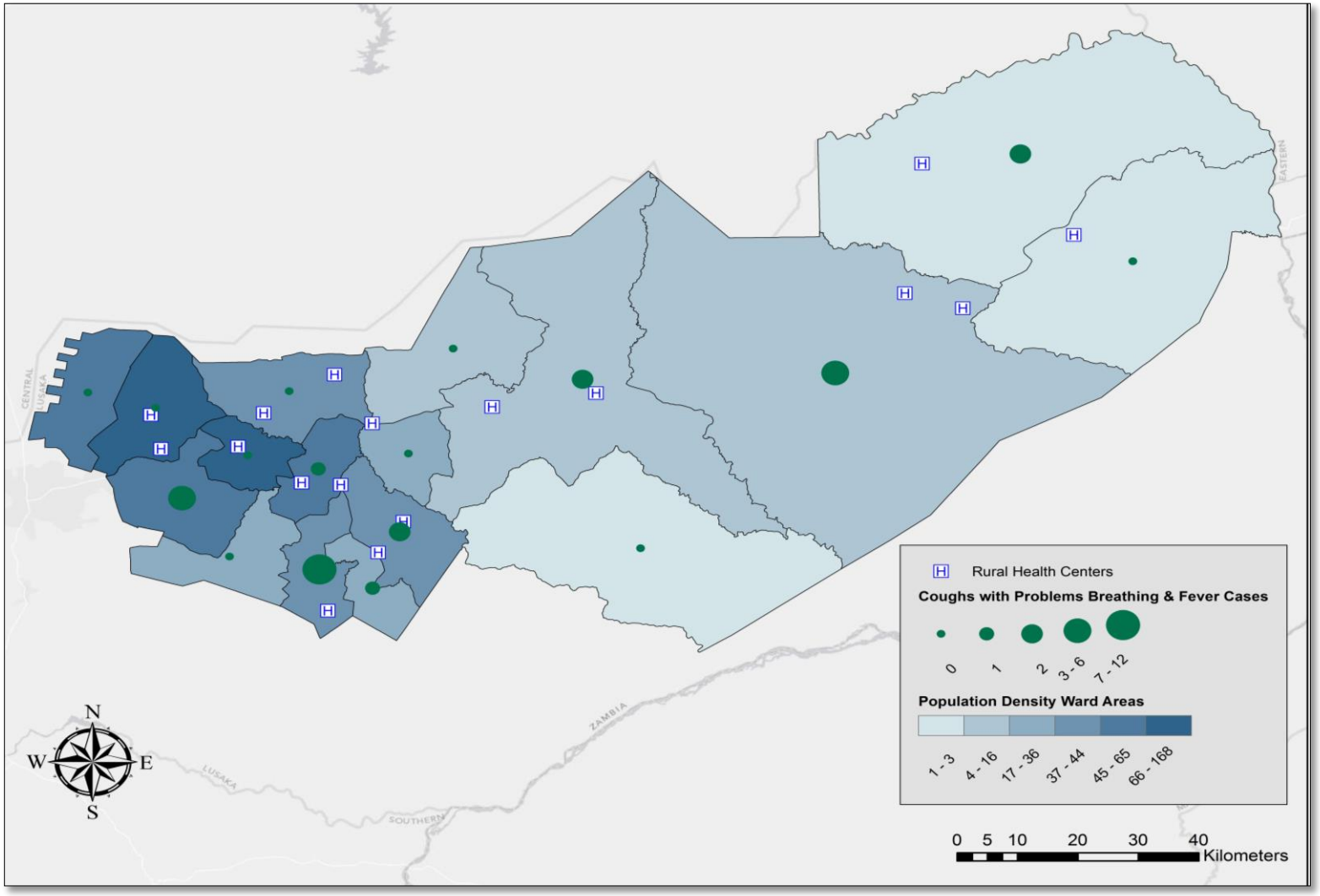
Cases of Fevers, Vomiting, & Diarrhea



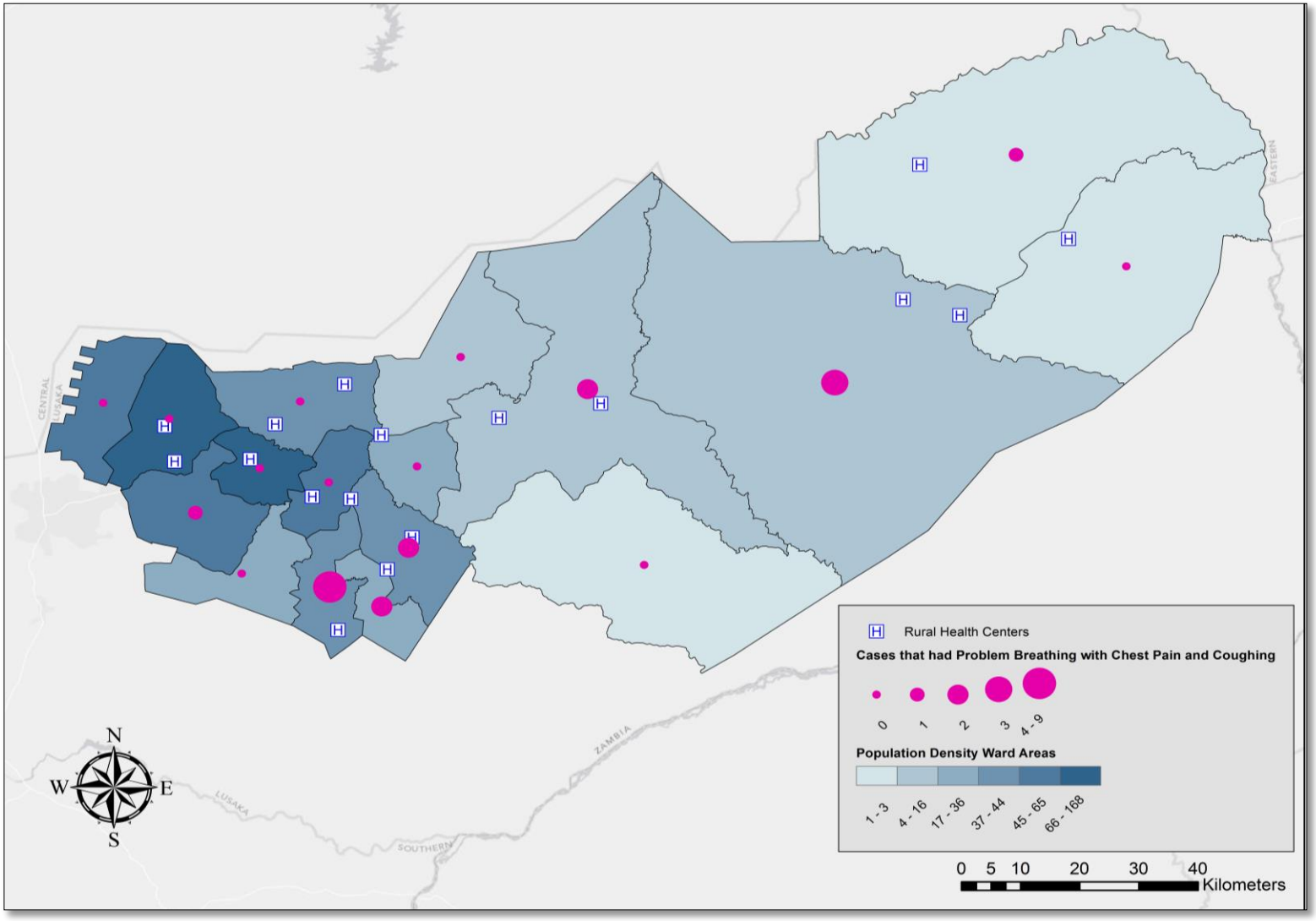
Cases of Malaria, Fever, and RDTs



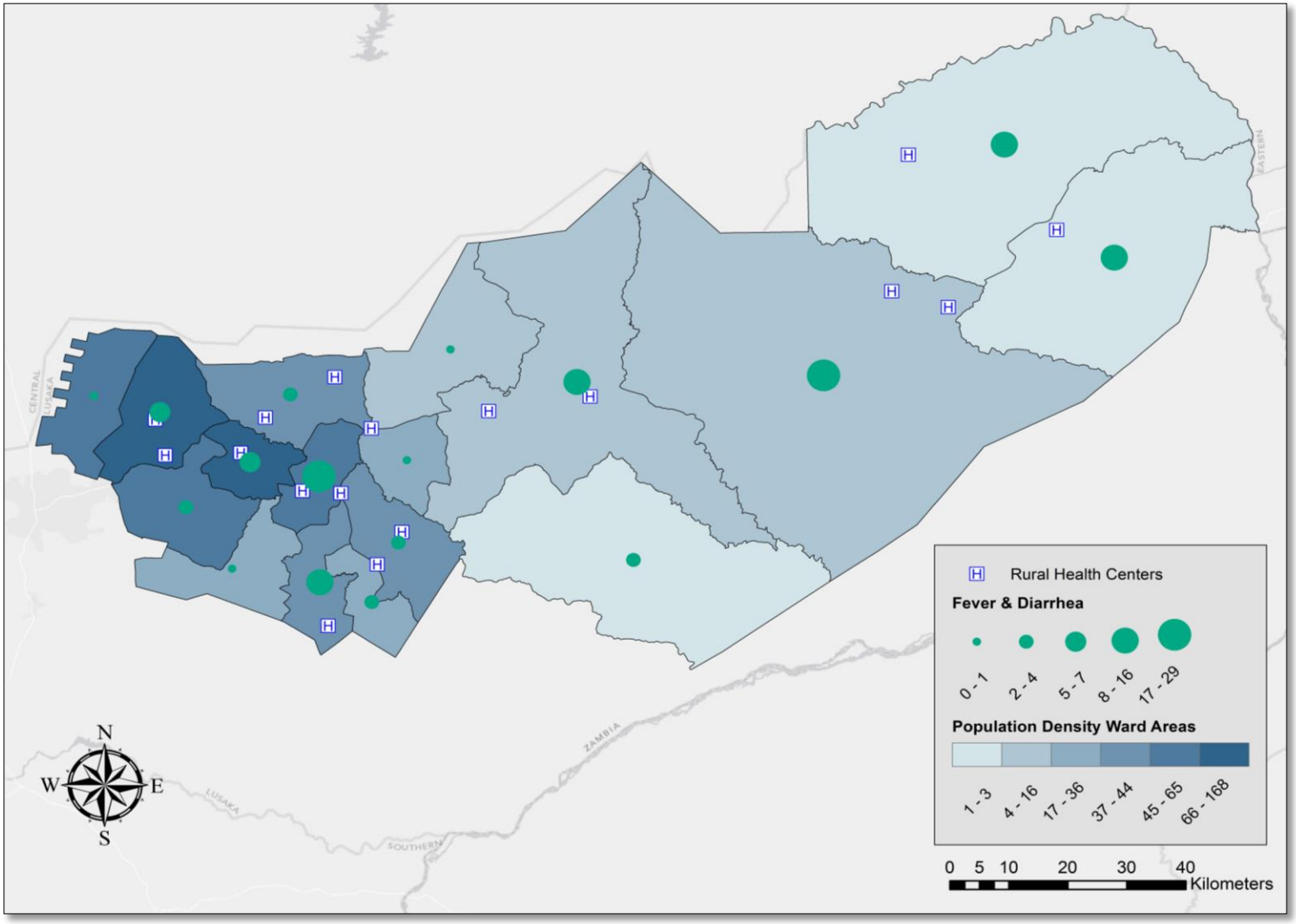
Cases of Pneumonia with Fever



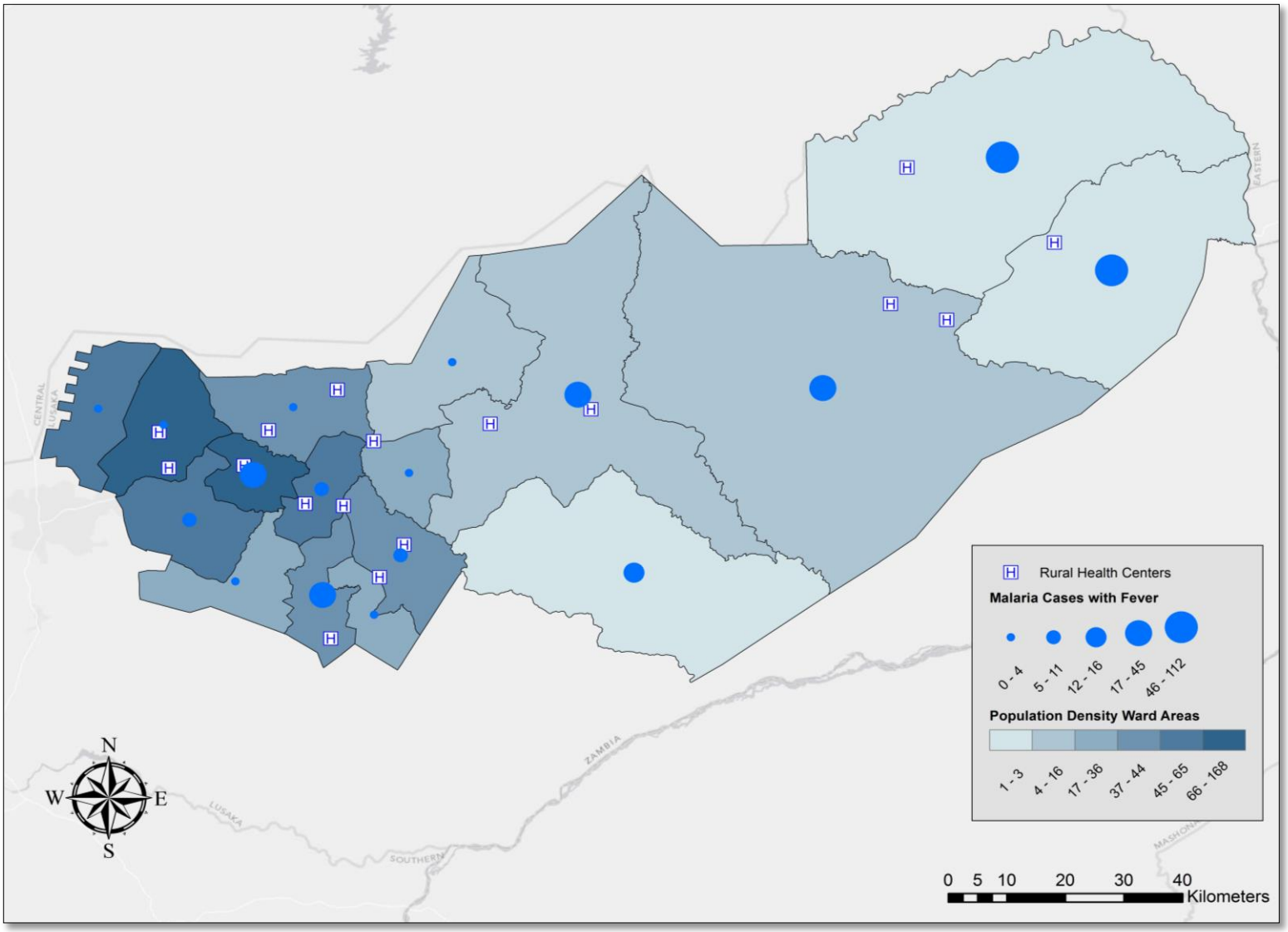
Cases of Problems Breathing, Chest Pain & Coughing



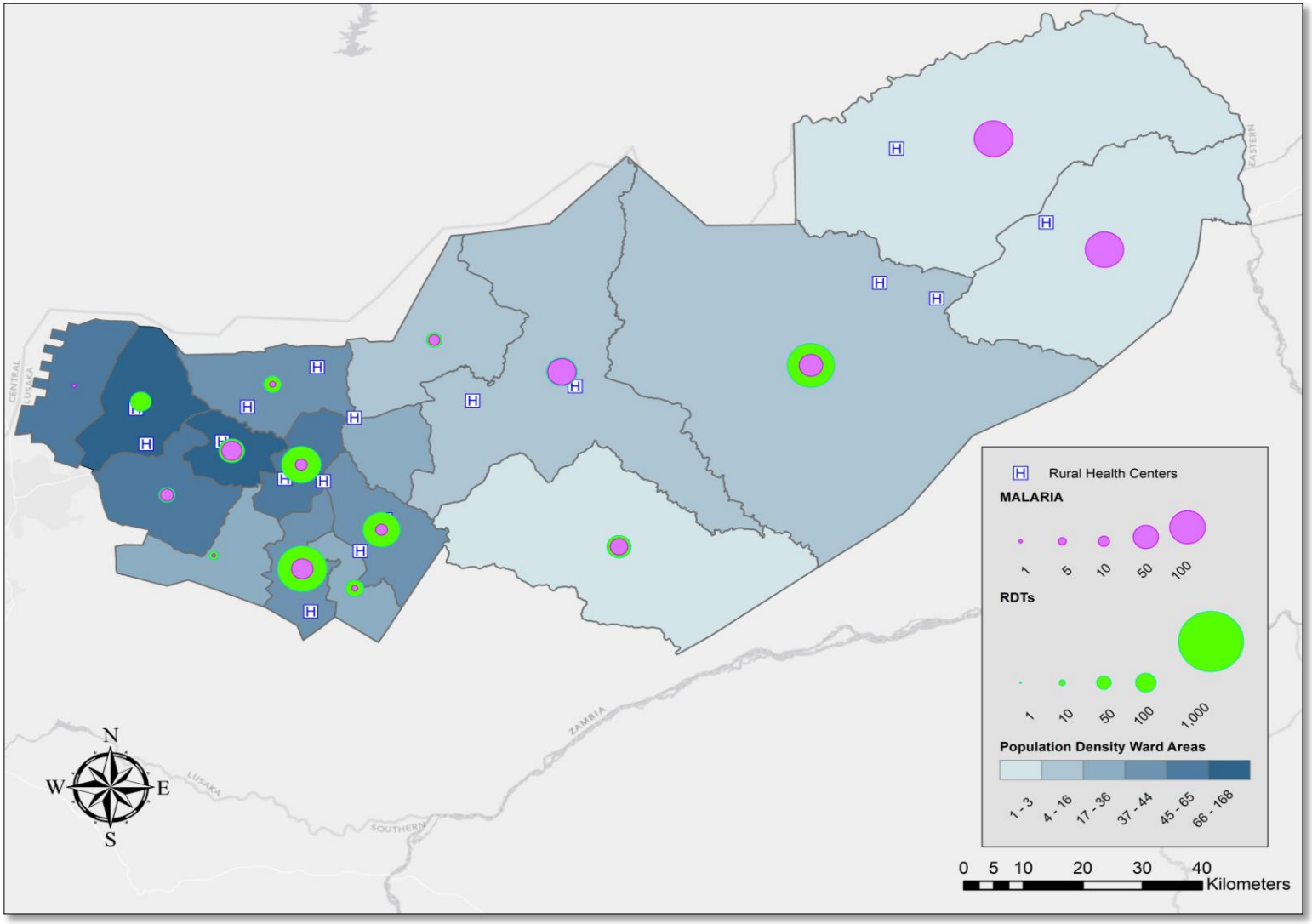
Cases of Fever and Diarrhea



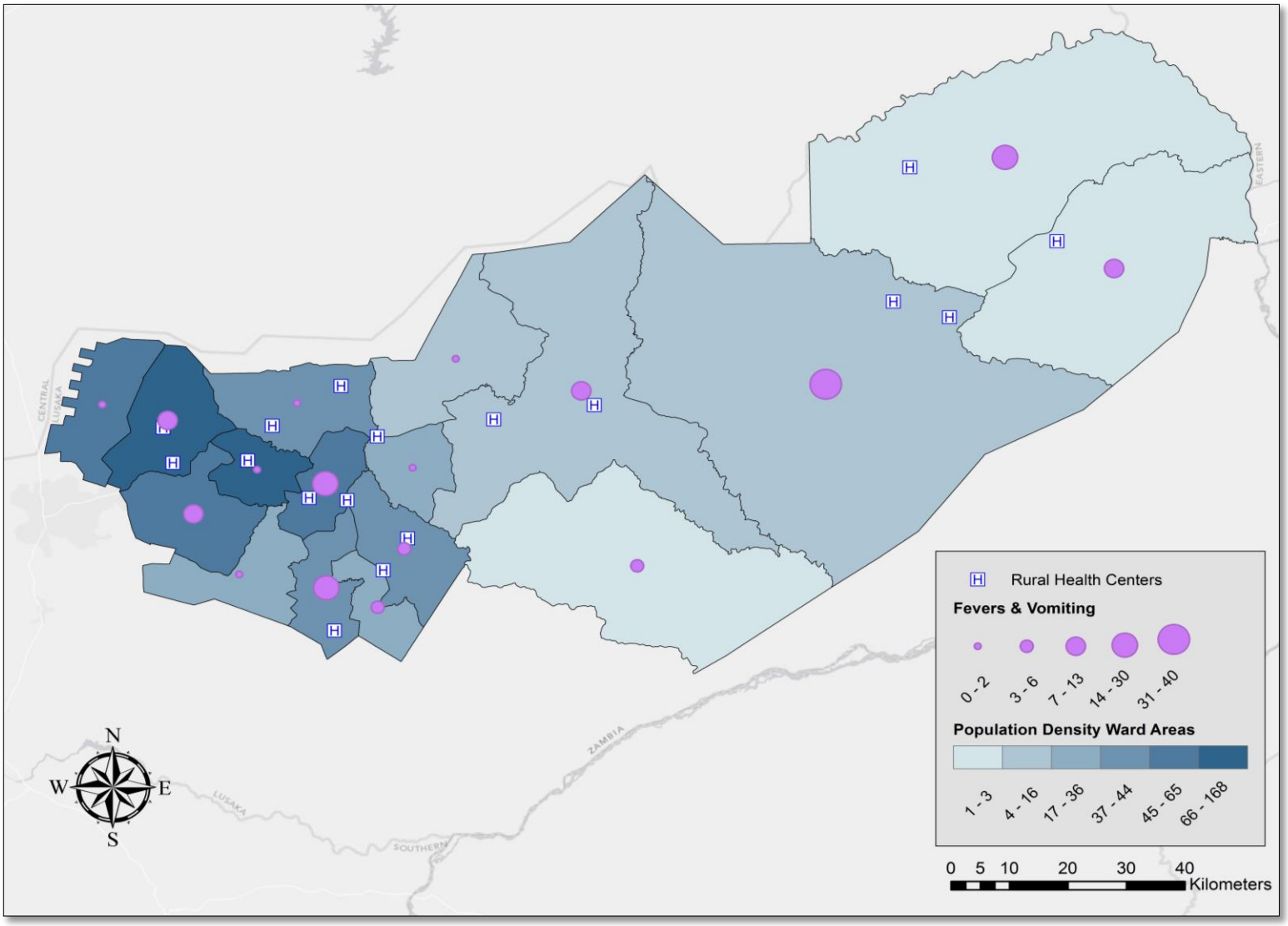
Cases of Malaria with Fever



Cases of Malaria and RDT Usage

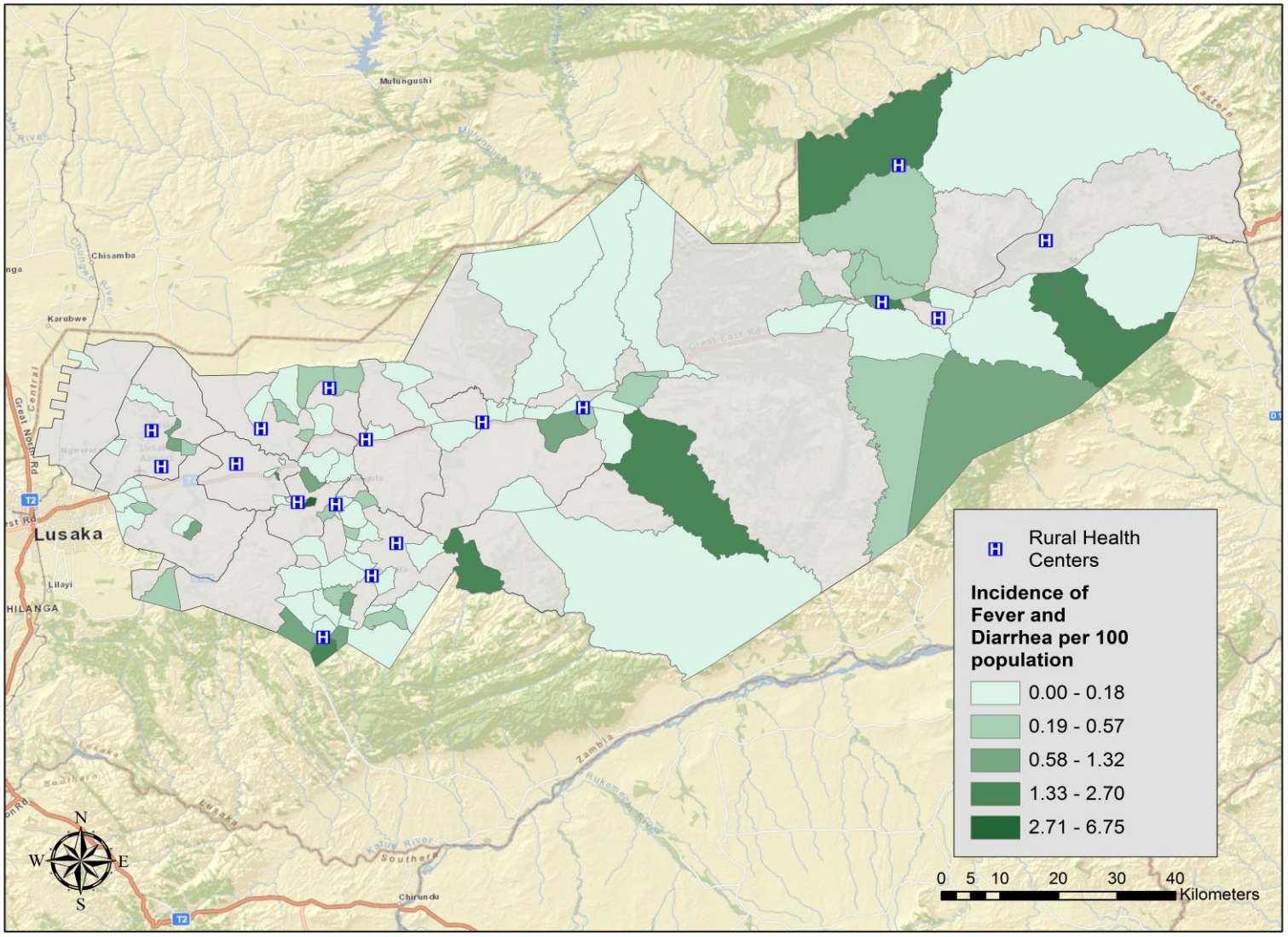


Cases of Fever & Vomit

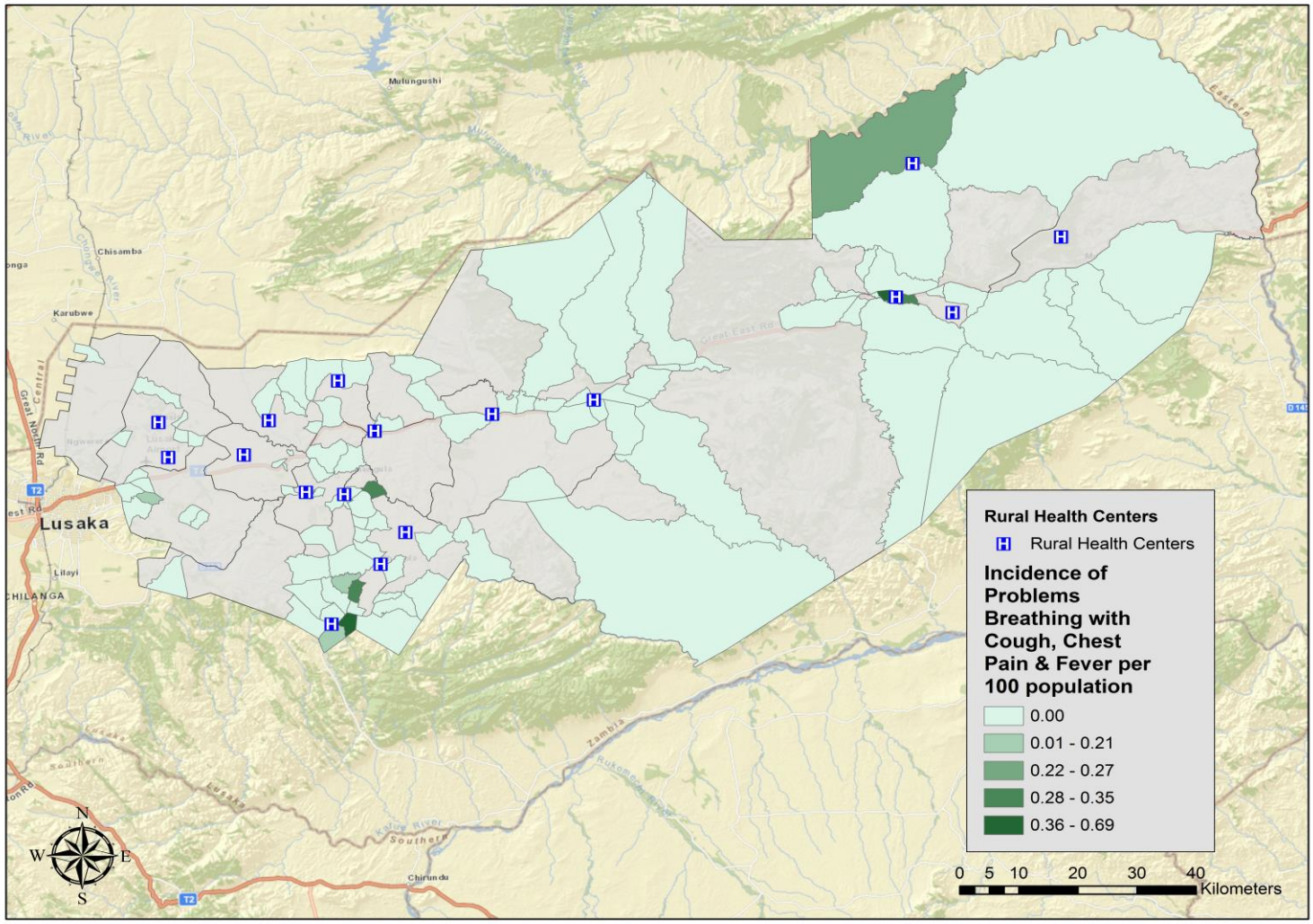


Appendix C: Disease/Illness Incidence Maps

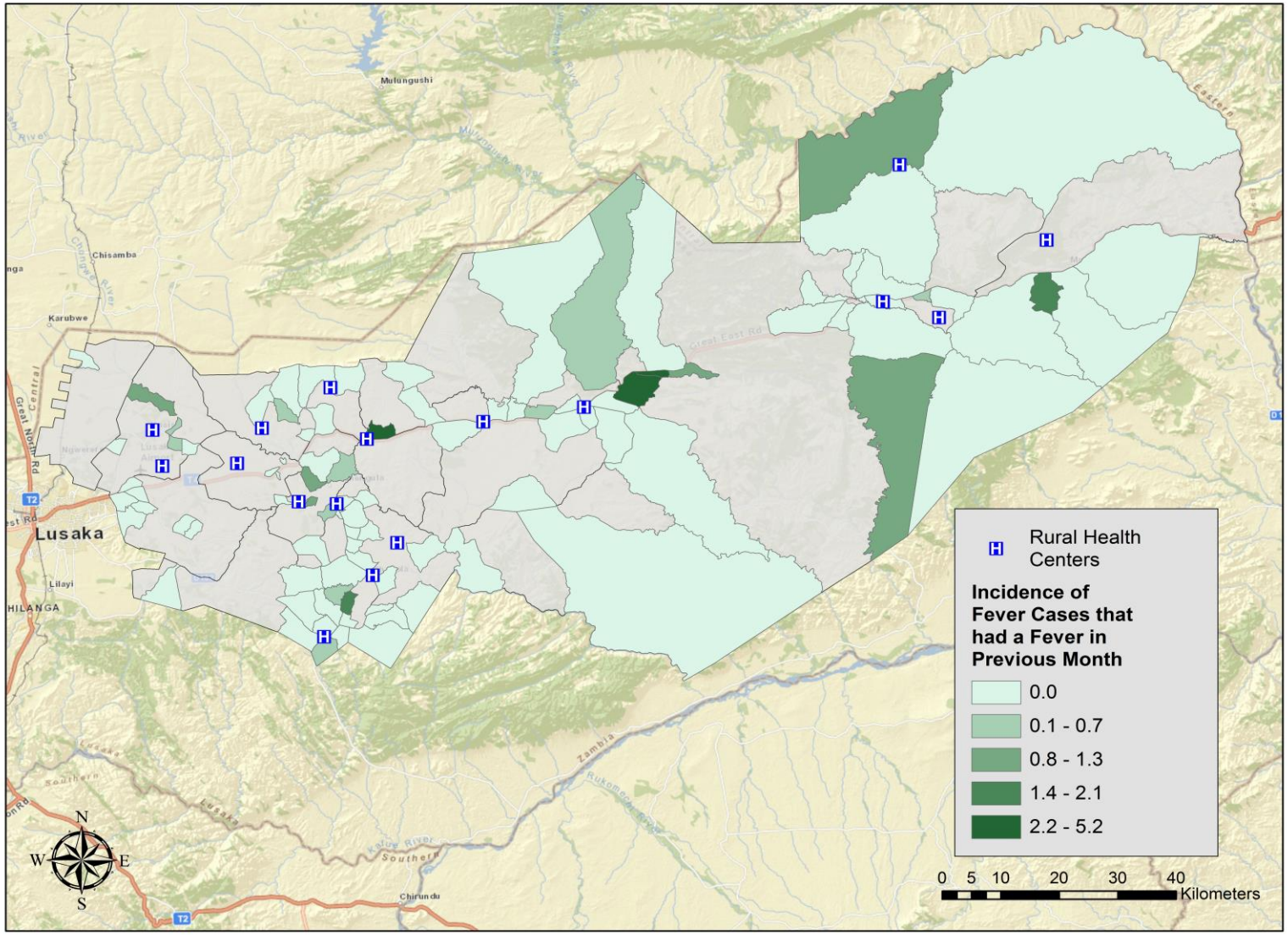
Incidence of Diarrhea & Fever



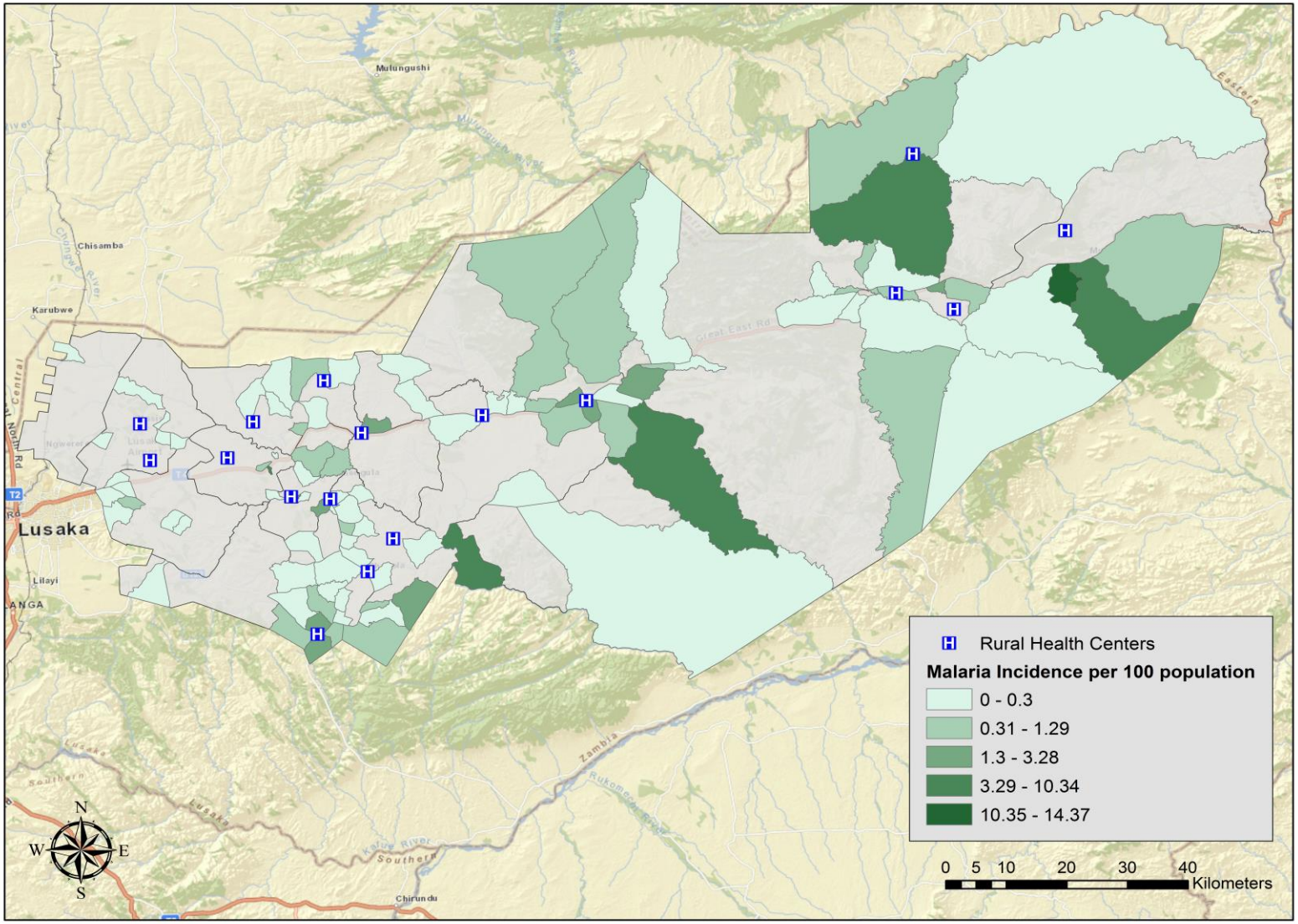
Incidence of Problems Breathing, Chest Pain, Cough and Fever



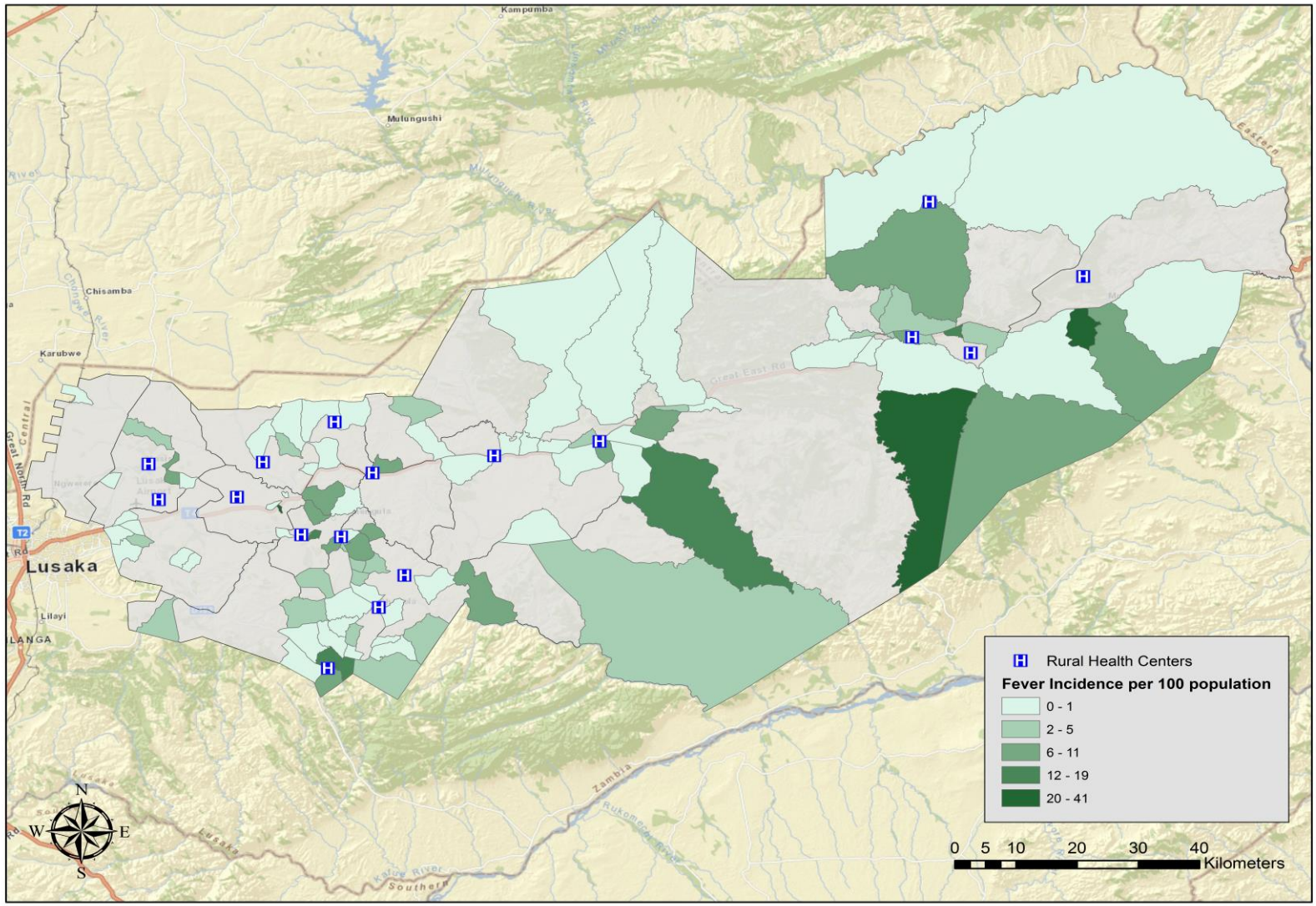
Incidence of Fevers with Fever History



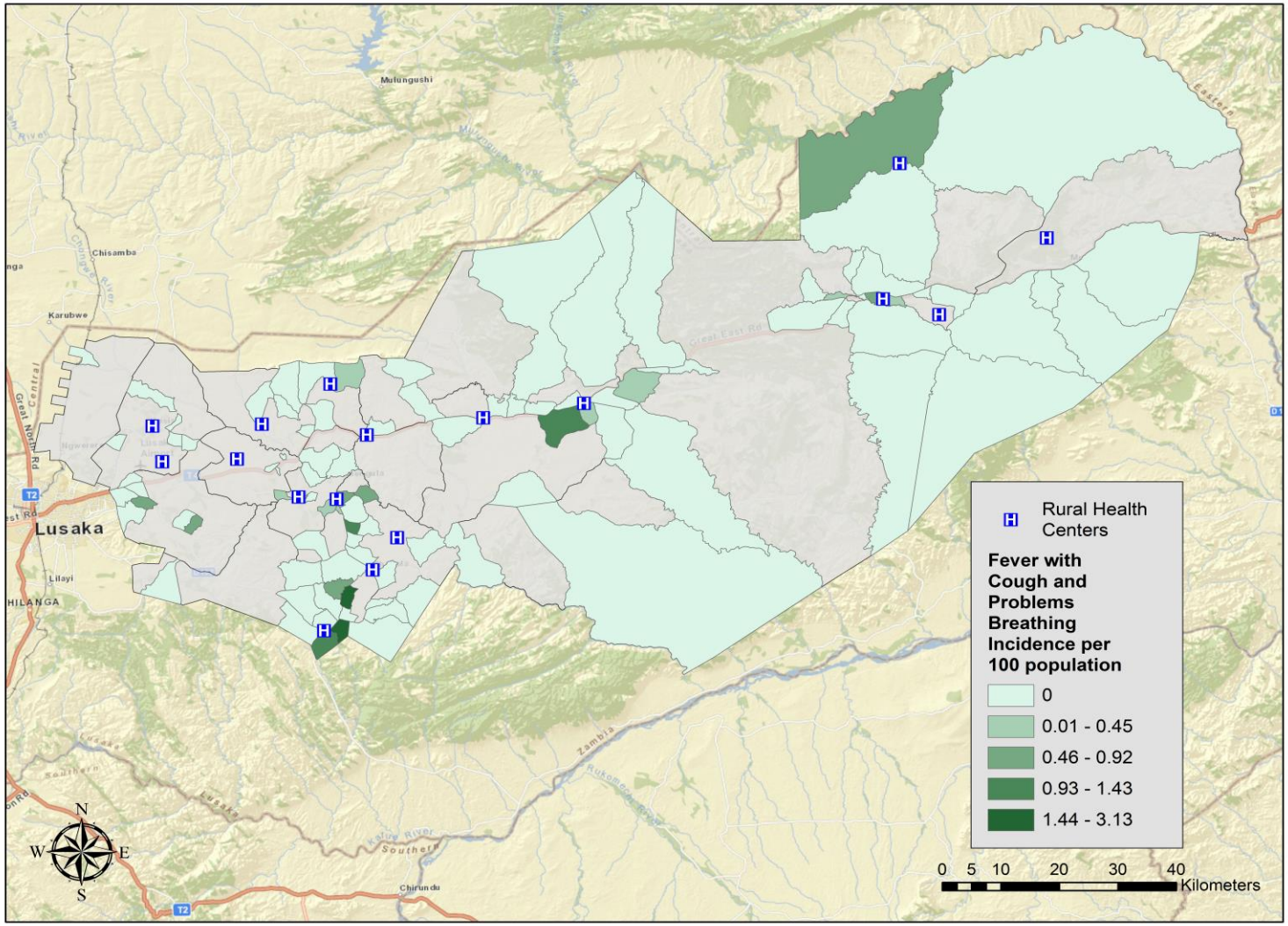
Malaria Incidence



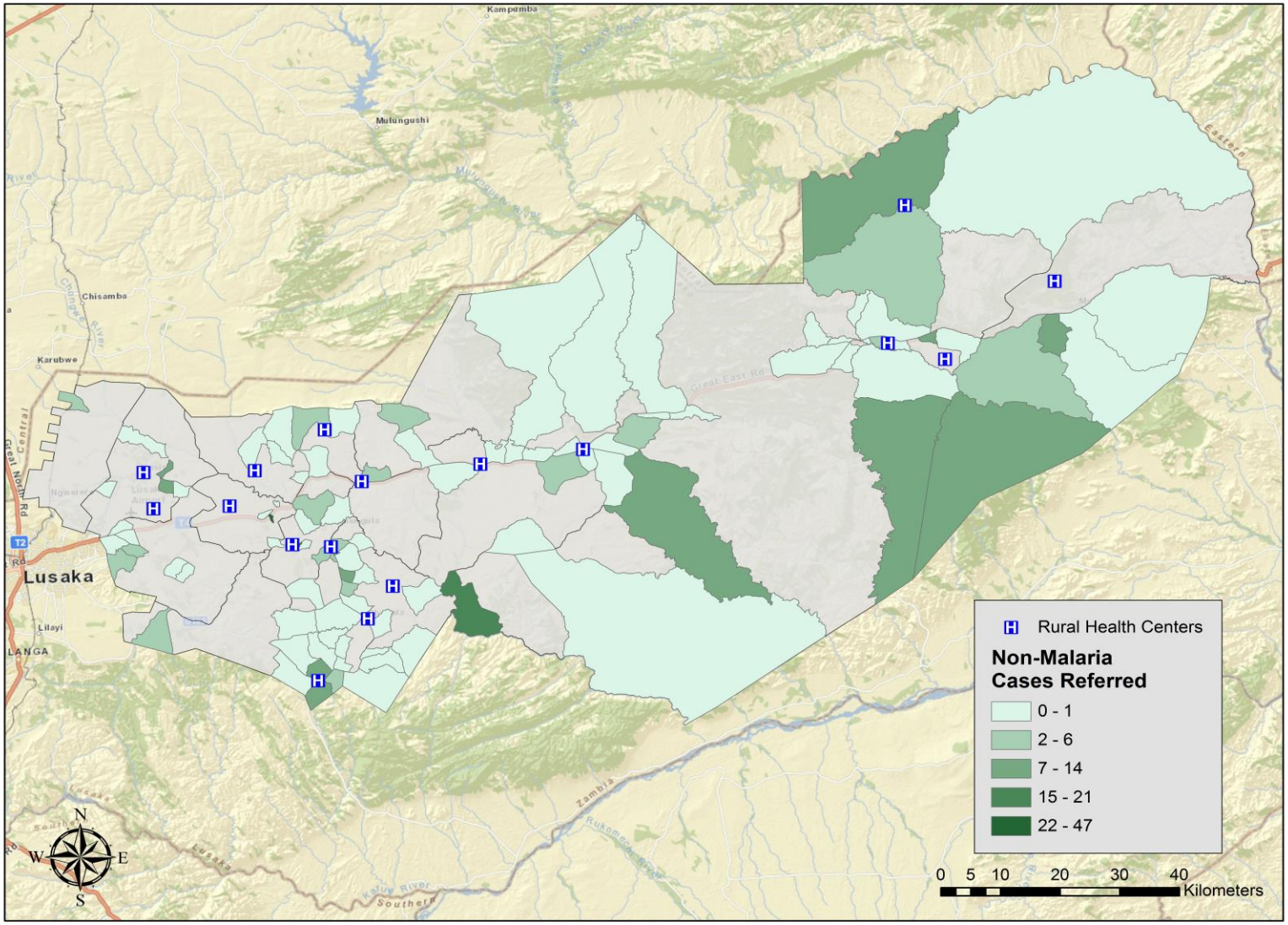
Incidence of Fever



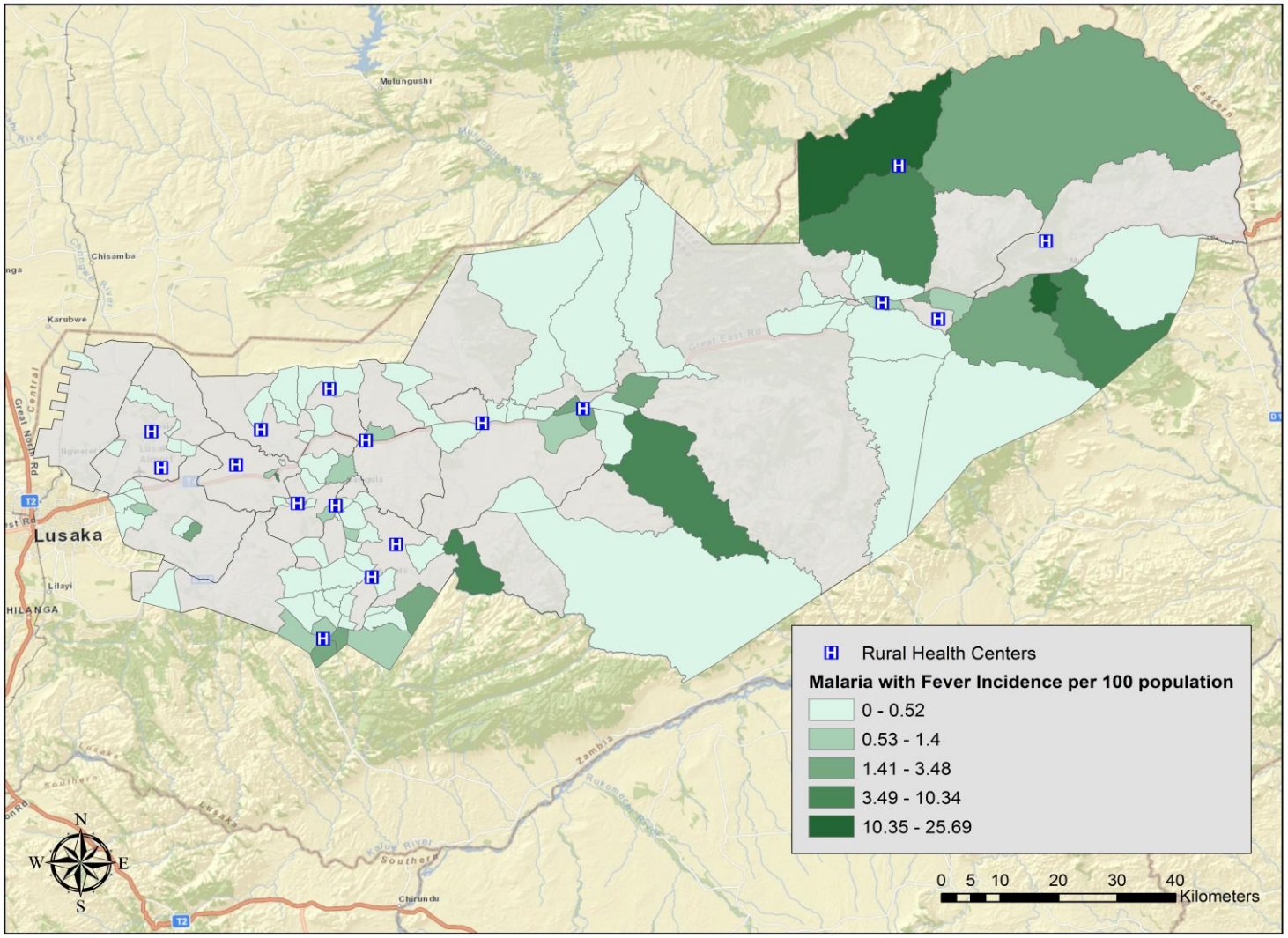
Incidence of Fever, Cough, & Problem Breathing



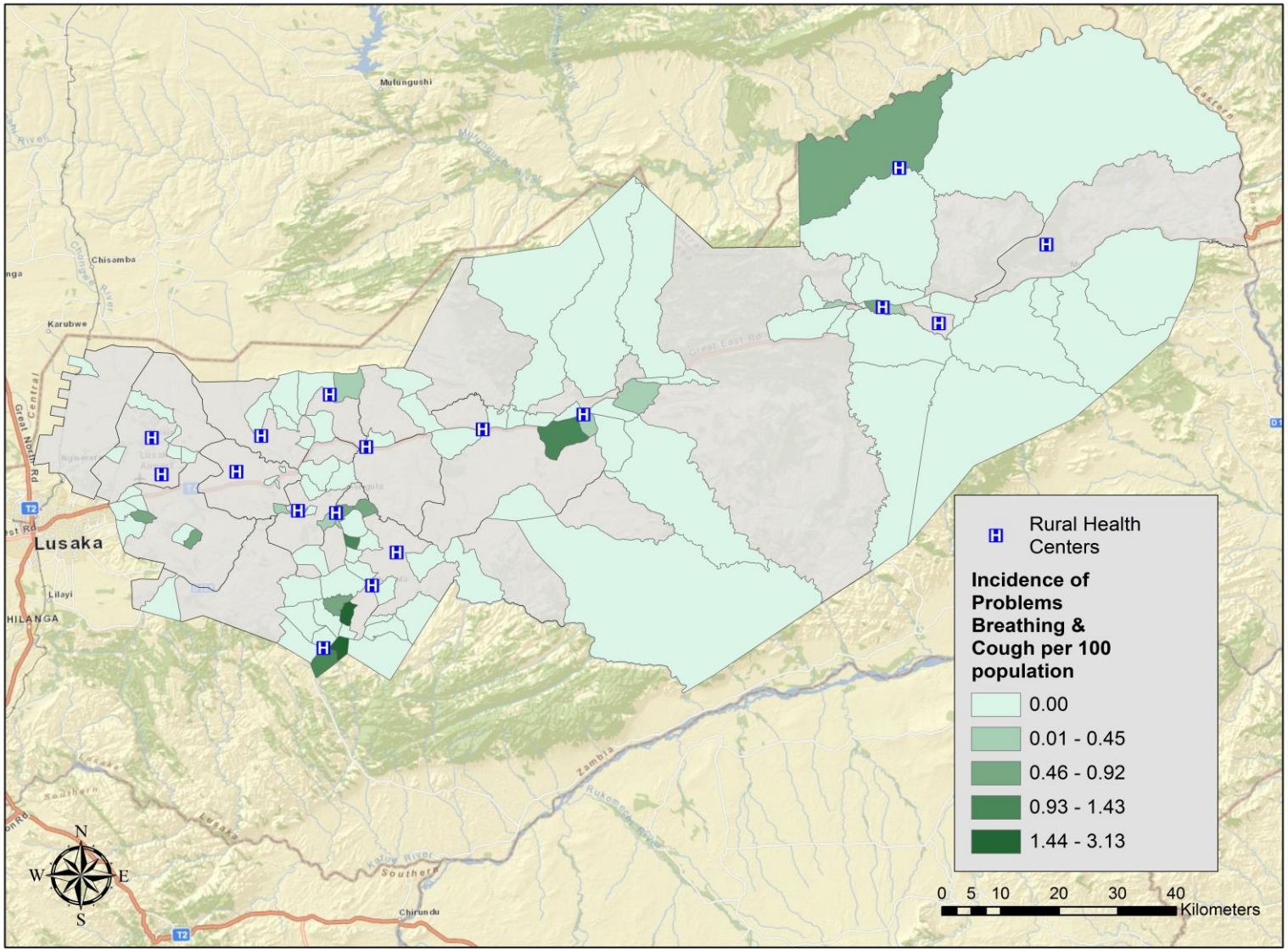
Incidence of Severe Diarrhea (Diarrhea & Vomiting) not referred



Incidence of Malaria with Fever



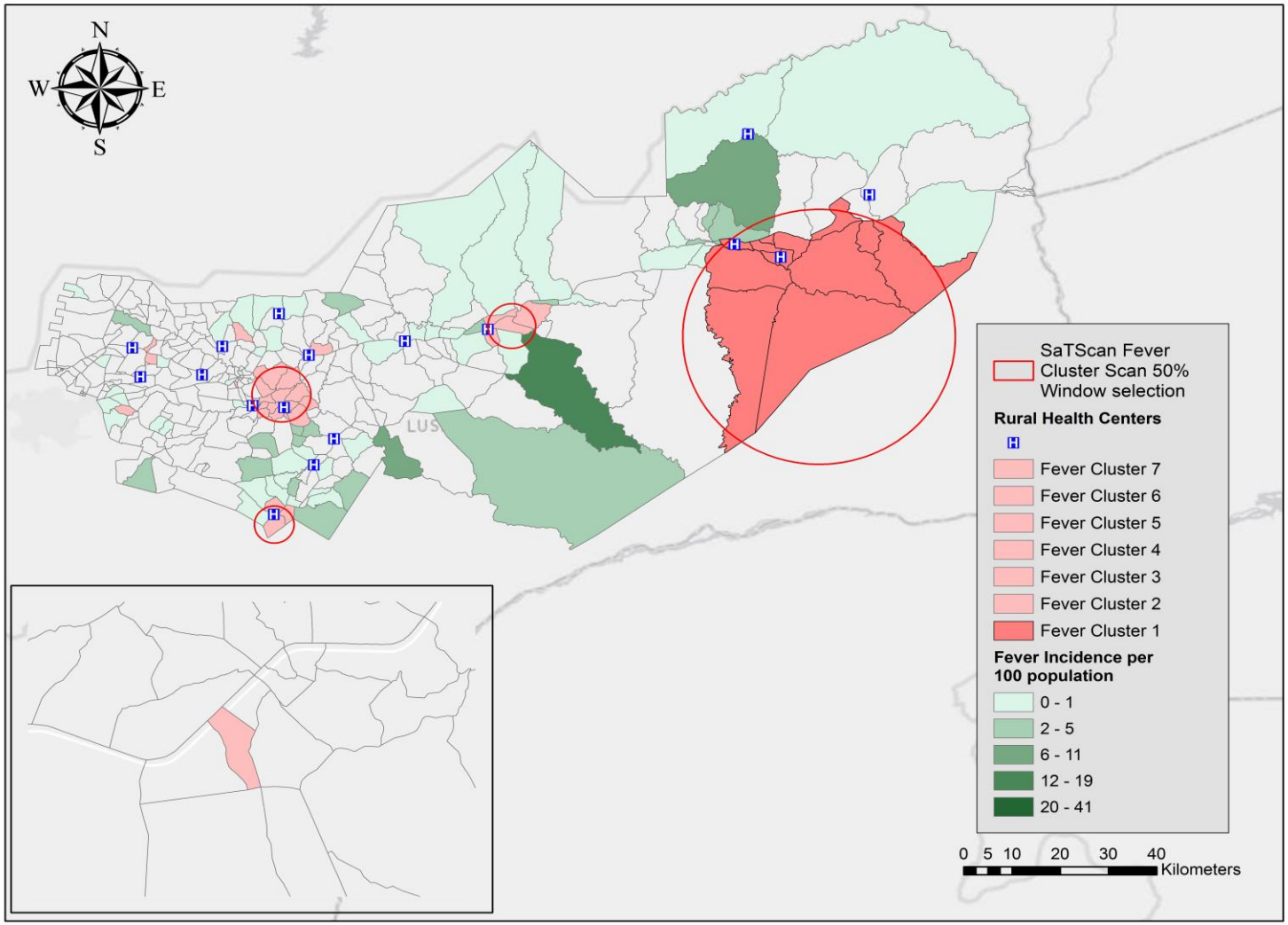
Incidence of Cough & Problems Breathing (Pneumonia)



Appendix D: SaTScan v9.3 Cluster Analysis Maps

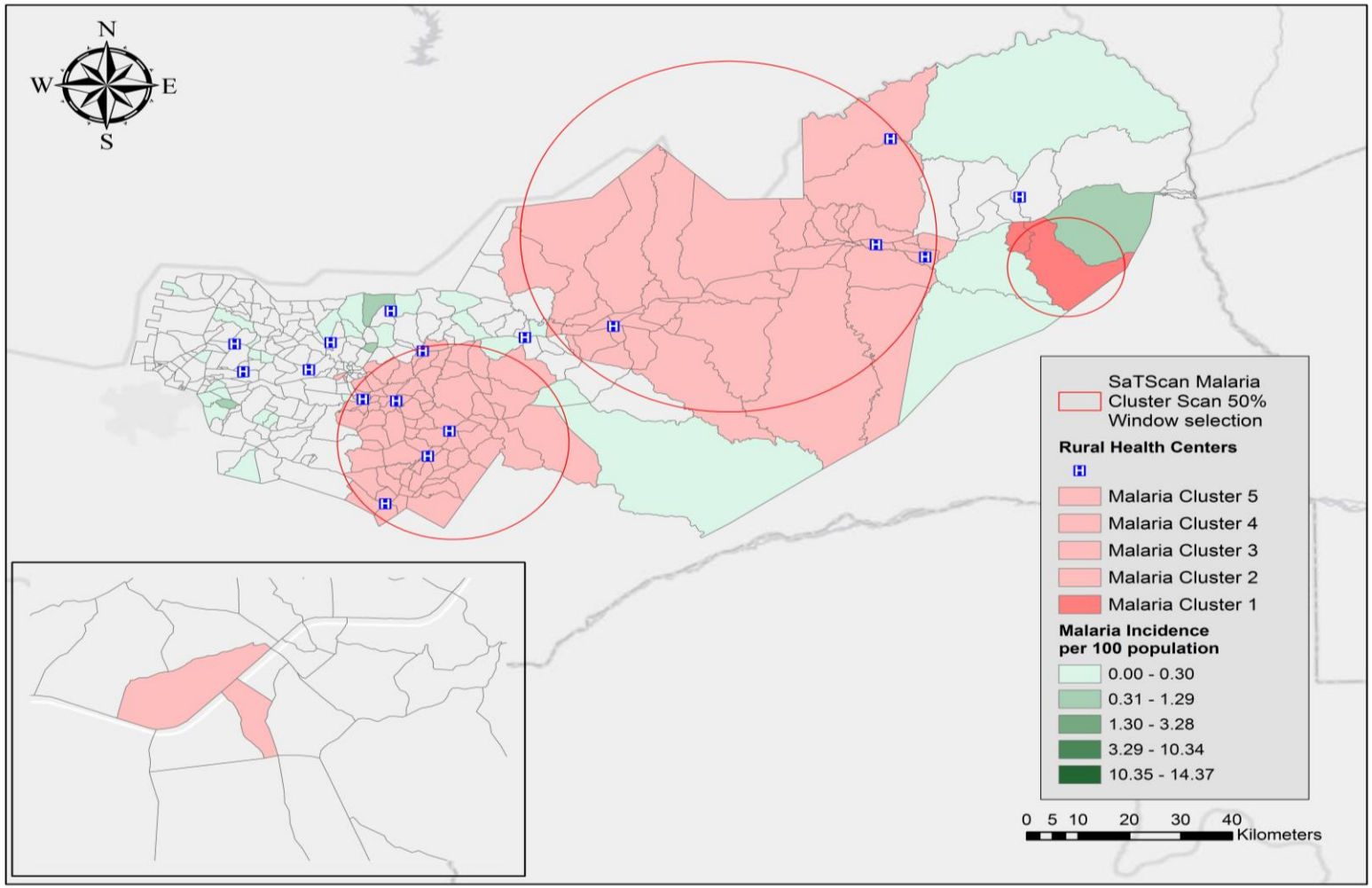
SaTScan v9.3 Fever Incidence Cluster Areas with 50% of Population at Risk for Fever in Chongwe SEA boundary areas

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



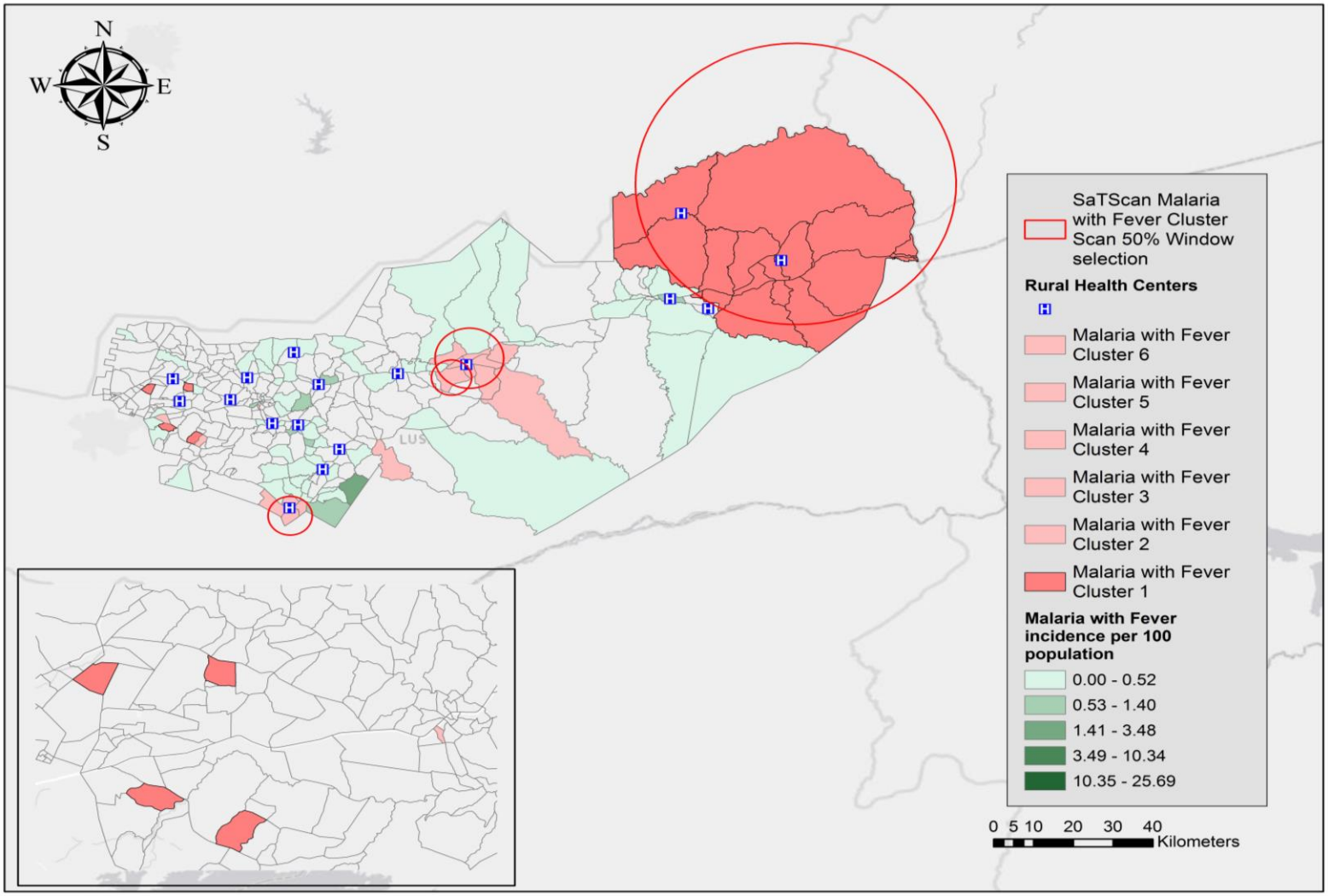
SaTScan v9.3 Malaria Incidence Cluster Areas with 50% of Population at Risk for Malaria in Chongwe SEA boundary areas.

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



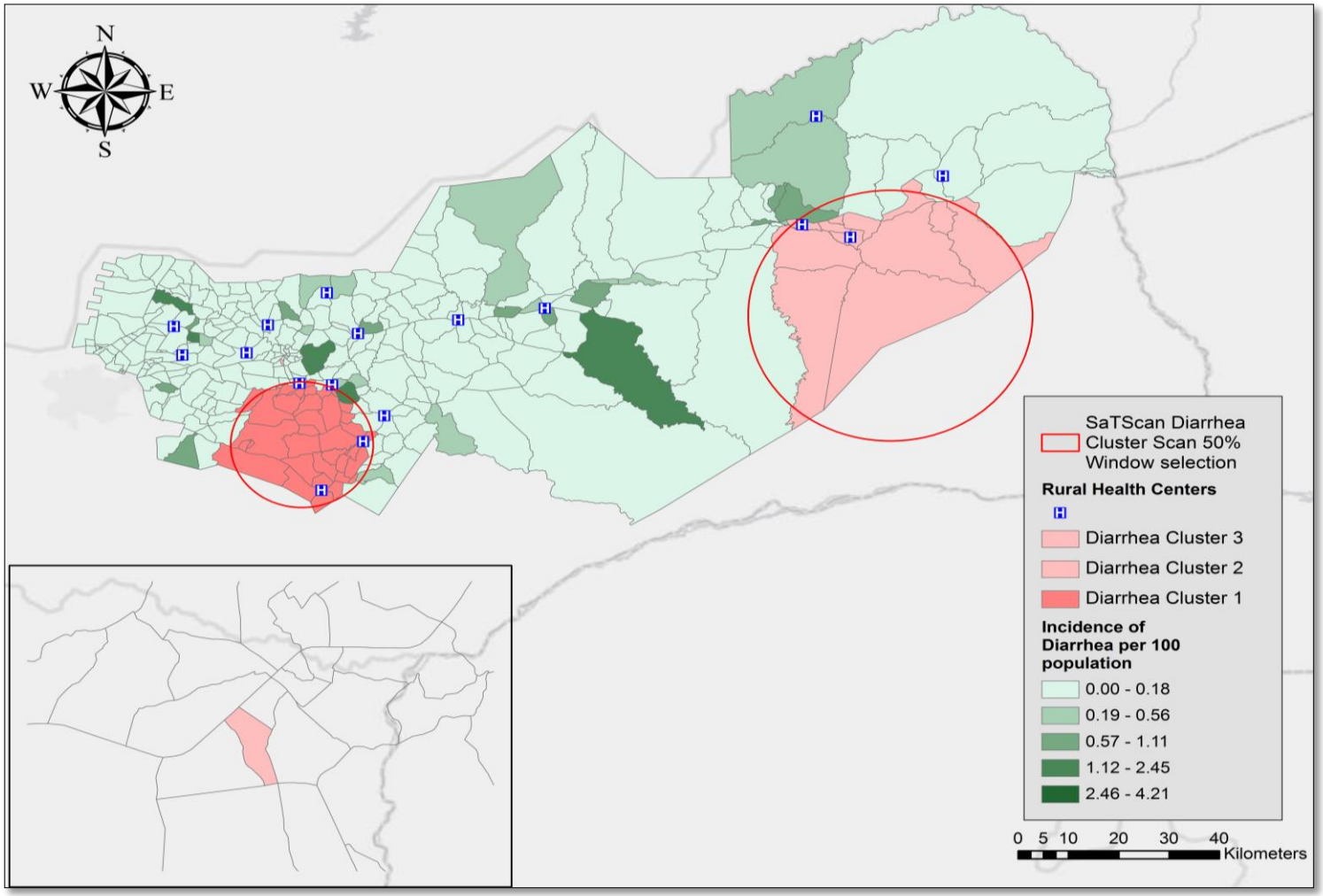
SaTScan v9.3 Malaria with Fever Incidence Cluster Areas with 50% of Population at Risk for Malaria with Fever in Chongwe SEA boundary areas.

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



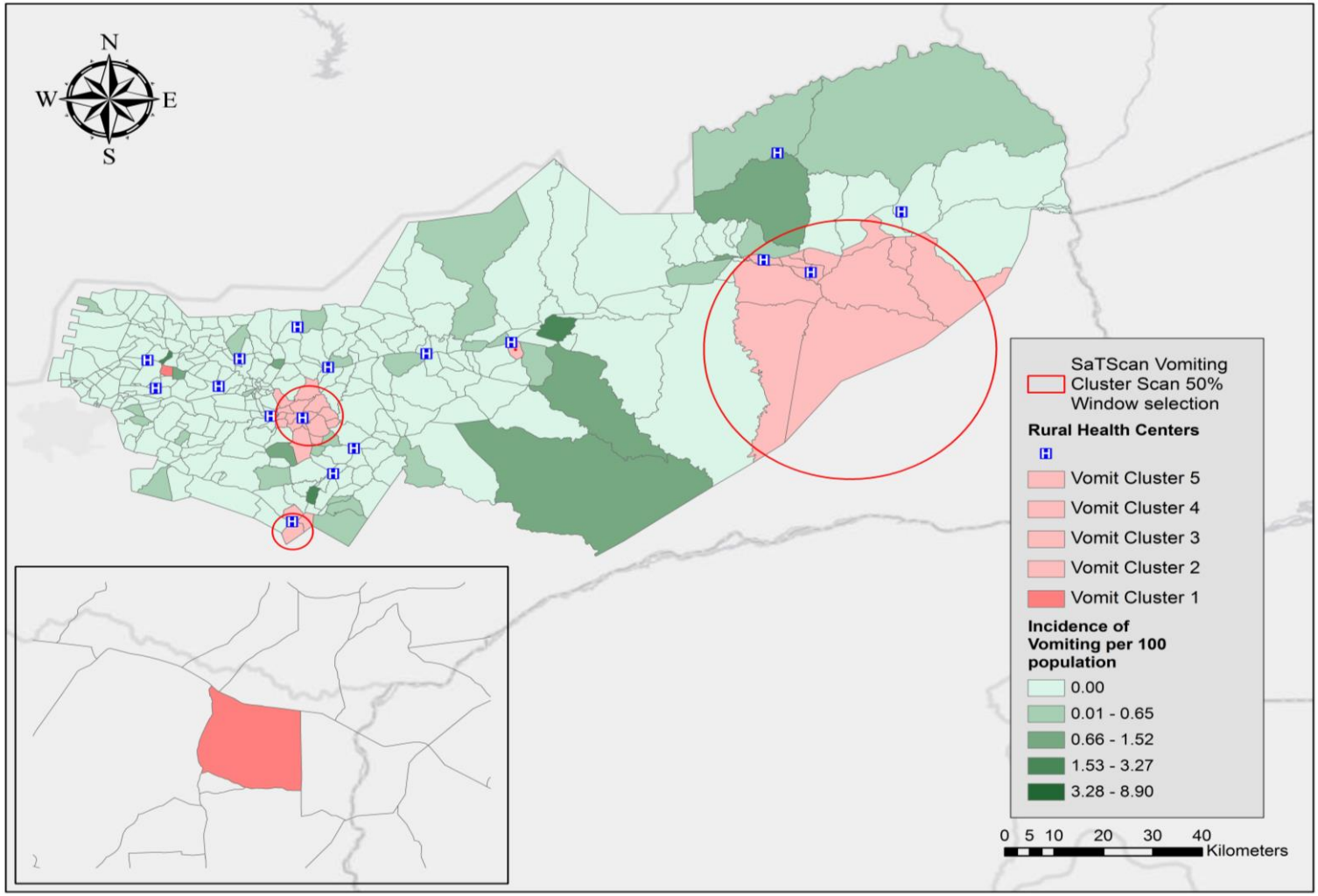
SaTScan v9.3 Diarrhea Incidence Cluster Areas with 50% of Population at Risk for Diarrhea Illness in Chongwe SEA boundary areas.

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



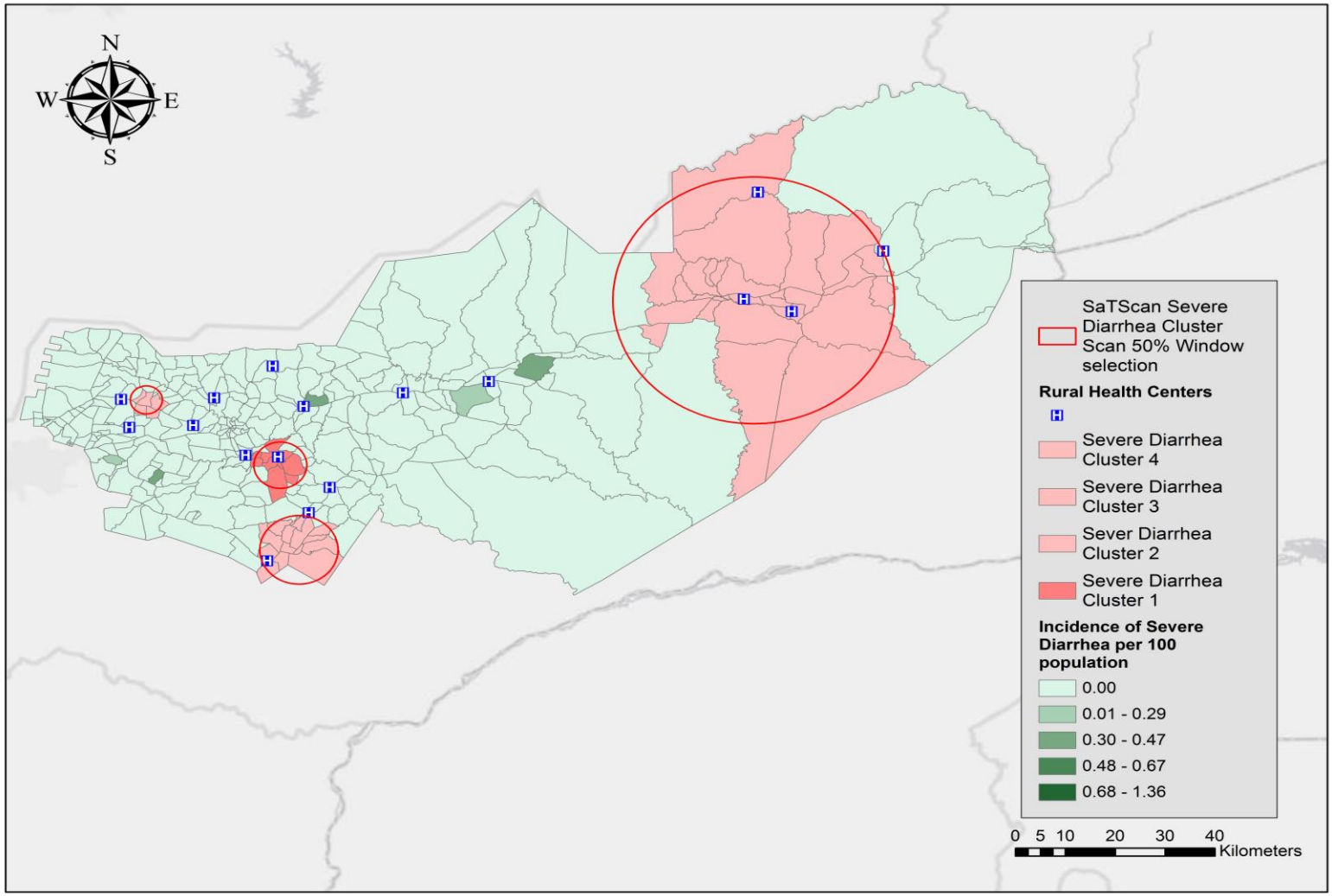
SaTScan v9.3 Vomiting Incidence Cluster Areas with 50% of Population at Risk for Vomiting Illness in Chongwe SEA boundary areas

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



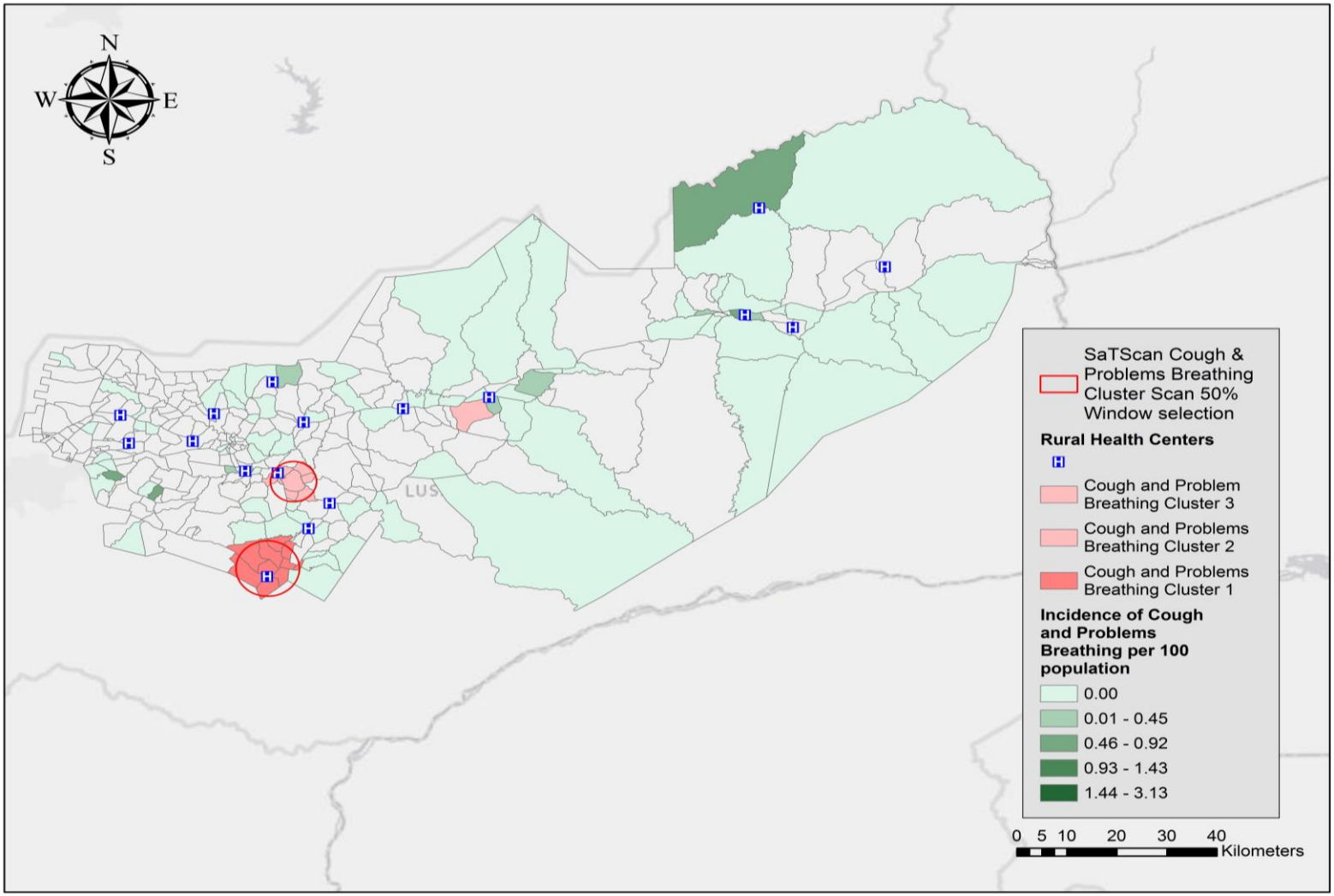
SaTScan v9.3 Severe Diarrhea Incidence Cluster Areas with 50% of Population at Risk for Severe Diarrhea Illness in Chongwe SEA boundary areas

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



SaTScan v9.3 Pneumonia Incidence Cluster Areas with 50% of Population at Risk for Pneumonia Illness in Chongwe SEA boundary areas

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**



SaTScan v9.3 Severe Pneumonia Incidence Cluster Areas with 50% of Population at Risk for Pneumonia Illness in Chongwe SEA boundary areas

***Most-likely (Primary) Clusters are marked in darker shades of red and Secondary Clusters in lighter shades of Pink**