CHIKUNGUNYA, DENGUE, AND ZIKA IN CALI, COLOMBIA: EPIDEMIOLOGICAL AND GEOSPATIAL ANALYSES

A dissertation submitted to
Kent State University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

by
Amy R Krystosik
December, 2016

Abstract

AMY R KRYSTOSIK, PH.D. CANDIDATE, DECEMBER, 2016

COLLEGE OF PUBLIC HEALTH, DEPARTMENT OF BIOSTATISTICS, ENVIRONMENTAL HEALTH SCIENCES, AND EPIDEMIOLOGY

CHIKUNGUNYA, DENGUE, AND ZIKA IN CALI, COLOMBIA: EPIDEMIOLOGICAL AND GEOSPATIAL ANALYSES

(120 PP.)

CO-DIRECTORS OF DISSERTATION: MARK A. JAMES AND MADHAV P. BHATTA

Dengue, chikungunya, and Zika are vector-borne diseases of global health concern. Cali, Colombia experienced hypoendemic dengue and outbreaks of Zika and chikungunya between October 2014 and April 2016. The epidemiological and geographical aspects of these diseases were investigated using classical epidemiological measures, geographically weighted regression modeling, and spatial video geonarratives. These diseases were found to be spatially clustered and related to key environmental, demographic, temporal, and climate variables at neighborhood and sub-neighborhood levels. These findings have implications for public health policy and vector control in Cali, Colombia.
Dissertation written by Amy R Krystosik
B.S., John Carroll University, 2010
M.P.H., Tulane University School of Tropical Medicine and Hygiene, 2011

Approved by
____________________________________, Co-Chair, Doctoral Dissertation Committee
Mark A. James, Ph.D., Professor and Chair – Biostatistics, Environmental Health Sciences, and
Epidemiology, Kent State University

____________________________________, Co-Chair, Doctoral Dissertation Committee
Madhav P. Bhatta, PhD, MPH, Associate Professor of Epidemiology, Kent State University

____________________________________, Member, Doctoral Dissertation Committee
Andrew Curtis, PhD, Co-Director of the GIS, Health & Hazards Lab, Professor, Department of
Geography, Kent State University

____________________________________, Member, Doctoral Dissertation Committee, Outside Discipline Person
Diana M. Dávalos, MD, MPH, DrPH, Head of Department of Public Health, Faculty of Health Sciences,
Icesi University, Cali, Colombia

____________________________________, Graduate Faculty Representative
Andrew Curtis, PhD, Co-Director of the GIS, Health & Hazards Lab, Professor, Department of
Geography, Kent State University

Accepted by
____________________________________, Department Chair
Mark A. James, Ph.D., Professor and Chair – Biostatistics, Environmental Health Sciences, and
Epidemiology, Kent State University

____________________________________, Dean, College of Public Health
Sonia Alemagno, Dean, College of Public Health, Professor Health Policy & Management Kent State
University
# TABLE OF CONTENTS

Chapter 1 ................................................................................................................................. 3

Background ............................................................................................................................. 3

Chikungunya ......................................................................................................................... 3

Dengue ................................................................................................................................. 3

Zika ......................................................................................................................................... 4

Control ................................................................................................................................... 4

The location ............................................................................................................................ 6

Methods ................................................................................................................................... 6

Ethics ....................................................................................................................................... 11

Significance and Statement of goals ..................................................................................... 12

Impact ..................................................................................................................................... 12

Chapter 2: ................................................................................................................................ 14

Synopsis .................................................................................................................................. 14

Manuscript I: ........................................................................................................................... 14

Abstract (250 words) ............................................................................................................ 14

Background ............................................................................................................................ 15

Chikungunya ......................................................................................................................... 15

Dengue ................................................................................................................................... 15
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping disease risk</td>
<td>73</td>
</tr>
<tr>
<td>Results</td>
<td>80</td>
</tr>
<tr>
<td>Discussion</td>
<td>85</td>
</tr>
<tr>
<td>Conclusion</td>
<td>89</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>92</td>
</tr>
<tr>
<td>Overarching Discussion with Conclusions and Limitations</td>
<td>92</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>92</td>
</tr>
<tr>
<td>Impact</td>
<td>93</td>
</tr>
<tr>
<td>References</td>
<td>95</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>IRB approval letters</td>
<td>107</td>
</tr>
<tr>
<td>Participant consent community health worker interview</td>
<td>112</td>
</tr>
<tr>
<td>Participant consent for audio and video recording</td>
<td>115</td>
</tr>
<tr>
<td>Author Resume</td>
<td>116</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1: Weekly and quarterly incidence ................................................................. 29

Figure 2: A. Count of dengue cases by month-year. B. Count of Chikungunya cases by week. C. Count of Zika cases by week D. Comparison of chikungunya, dengue, and Zika cases by quarter. D Precipitation and temperature anomalies, Cali, Colombia .................................................................................. 30

Figure 3: a. dengue October 2014 - April, 2016, b. chikungunya January 2015 - April, 2016. C. Zika November 2015 - April, 2016. Click here http://capturadedatos.co/surveillance/# to view the evolution of these diseases in Cali, Colombia. login id: dengue, Password: losmapasdecali .......................................................... 31

Figure 4: Weighted incidence per 100,000 population by neighborhood. a. Zika November 2015 - April, 2016, b. dengue October 2014 - April, 2016, c. chikungunya January 2015 - April, 2016 .................................................. 31

Figure 5: Incidence Ratio (Female: Male). A. By outcome and year. B by age category and year. Zika November 2015 - April, 2016, dengue October 2014 - April, 2016, chikungunya January 2015 - April, 2016. .............................................................................. 33

Figure 6: Distribution of days from onset of symptom to exam by test result. ................................................. 34

Figure 7: A. Count of dengue cases by month-year. B. Count of Chikungunya cases by week. C. Count of Zika cases by week; D. Average rain (mm) and temperature anomalies (°C) from October, 2014 (10) to April, 2016 Rain data (Kuligowski 2016); Temp anomalies data (Morice 2012) ......................................................... 42

Figure 8: Risks by neighborhood ...................................................................................................... 47

Figure 9: Chikungunya coefficients and standard error, January 2015 - April, 2016 ........................................ 51

Figure 10: Dengue coefficients and standard error, October 2014 - April, 2016 ................................................. 52

Figure 11: Zika coefficients and standard error, November 2015 - April, 2016 ..................................................... 53
Figure 12: small drainage canal ..............................................................67
Figure 13 a and b: Two Contour 2 cameras mounted on the inside left and right window of a vehicle-angled down to capture the conditions by the side of a road); C: hand recording of mosquito larvae by the two individuals seen in Figure B ........................................................................................................70
Figure 16: Word cloud of geospatial interview transcripts. ..............................................................................................................77
Figure 17: SVG routes ..........................................................................................................................................................79
Figure 18: mapping of the Calipso neighborhood; KDE of spatial risk classified into ten equal categories and contoured ........................................................................................................................................81
Figure 19: Google street view versus Spatial video: KDE of water classified into ten equal categories and contoured ........................................................................................................................................82
Figure 20: KDE of trash and tires classified into ten equal categories and contoured .................................................................84
Figure 21: Top 50% Contours of Risk by Disease. Surveillance data for Dengue, Chikungunya, and Zika are overlaid on the path buffer. Dengue October 2014 - April, 2016. Zika November 2015 - April, 2016. Chikungunya January 2015 - April, 2016 ........................................................................................................................................86
Figure 22: small pocket parks with playsets overlaps point “E” .....................................................................................................88
LIST OF TABLES

Table 1: Sample characteristics frequency and percentage by outcome ................................................. 27
Table 2: Census 2005 characteristics by neighborhood ............................................................................. 39
Table 3: Series of models for each disease outcome .................................................................................. 44
Table 4: AIC comparing global and GWR Poisson models Chikungunya .............................................. 48
Table 5: AIC comparing global and GWR Poisson models Dengue .......................................................... 49
Table 6: AIC comparing global and GWR Poisson models Zika .............................................................. 50
Table 7: comparison of models.................................................................................................................. 54
Table 8: spatial video and spatial video geonarrative field collections and resulting mapped length, words, and spatial mentions .......................................................................................... 78
**Preface**

This body of work is the result of two years of collaboration and learning. The original idea for the project came a need to better understand the chikungunya outbreak that was occurring during my time living in Cali, Colombia. The resulting project is a compilation of the expertise of many mentors and co-workers who became close personal friends and colleagues.

We acknowledge the contributions of SIVIGILA, The Secretary of Health of Cali Colombia, the community health workers and community mothers of Cali, ICESI University, The Geography Lab at Kent State University, Valle de Lili Hospital, and Cañaveralejos Hospital in Cali, Colombia for their contributions to this work.
Chapter 1

Introduction

Background

Zika \(^1\), dengue \(^2, 3, 4, 5\) and chikungunya \(^6, 7\) are vector borne diseases transmitted by bites of an infected *Aedes aegypti* mosquito. Zika can also be transmitted through sexual contact \(^8, 9, 10\). Individually, each represents an important, global public health threat.

Chikungunya

An estimate of 1.3 billion persons live in areas where chikungunya can be transmitted \(^11\). Chikungunya was responsible for 1,386 - 1,081,962 years of life lost in 2005 \(^12\). Chikungunya fever is caused by an alphavirus and characterized by fever and severe joint pain. Chronic arthritis often results but has not yet been consistently defined \(^13, 14, 15\). There is increasing global concern over expanding distribution \(^16, 17, 18, 19, 20, 21\). Since introduction in 2013, CDC estimates more than 1.7 million suspected cases in 45 countries in the Americas \(^19\). However, there is confusion between probable and confirmed cases \(^22\) as the disease is often co-endemic and confused with dengue fever \(^18, 23, 24\) and Zika \(^25\).

Dengue

Dengue is responsible for 390 million infections per year \(^26\). Dengue fever is caused by one of four dengue virus (dengue) serotypes from the family *Flaviviridae* and is characterized by fever, rash, headache, and body aches and in severe cases, shock and hemmorhage. Severity is classified as dengue without warning signs, dengue with warning signs, and severe dengue \(^27\). The *Aedes aegypti* mosquito is
known to be the main vector of the virus, with *Aedes albopictus* transmitting another ~100 million cases per year.

**Zika**

Zika is a reemerging vector-borne viral disease declared to be a public health emergency of international concern in 2016 by WHO Director General, Margaret Chan. Zika was reclassified as a continued threat November 18, 2016. Zika virus is a flavivirus known to circulate in Africa, the Americas, Asia, and the Pacific with recent outbreaks in the Americas. Zika causes acute febrile illness and is associated with microcephaly and other neurological disorders in infants with in-utero exposure. It is transmitted by the same vectors as dengue and chikungunya, *Aedes albopictus* and *Ae. aegypti*. Twenty five percent of infected individuals exhibit symptoms lasting 2-7 days. Symptoms include mild fever, rash, conjunctivitis, and muscle pain. In epidemic areas, an increased number of neurological syndromes have been observed including Guillain-Barré Syndrome, Guillain–Barré Syndrome, meningoencephalitis, and myelitis. Zika has also been associated with congenital microcephaly and other birth defects, with increased incidence of microcephaly observed in children born to pregnant women infected with Zika. The burden of Zika falls on the poor since they cannot afford to limit the risk factors of this disease, which include conditioning, window screens, or even insect repellents.

**Control**

Despite recent increases in reported outbreaks and distribution, no vaccines or cure for these diseases exist; however, new vaccines are being developed. The Zika virus vaccines are showing promise with plans for clinical trials and some nucleoside analogue drugs have shown promising cellular activity. Chikungunya has no licensed vaccines but currently holds around 15 candidates that are in development and several others are undergoing clinical trials. The first dengue vaccine was licensed for those age 9-45 years of age living in endemic areas, and two more are beginning phase three of
investigation in 2016. No specific treatment exists for any of these diseases, rather the focus is on relieving the patient's symptoms.

Vector control for the Aedes mosquito vector remains the primary prevention method. Common risk factors for dengue, chikungunya, and Zika are environmental as the Aedes mosquito can breed in any container of stagnant water (e.g. flower pots, puddles in road, tree holes, tires, bromeliads, buckets, drums, small trash, bottles, cans) and are well adapted to urban environments. Traditional methods of vector control include indoor residual sprays, fumigation, and source reduction. Biological forms of vector control are gaining popularity. For example, Sterile Insect Techniques releases sterilized mosquitoes, to reduce the population by means of failed reproductive events. In Medellin, Colombia, Wolbachia infected Ae. aegypti are being released and have reduced vector competence for ZIKV. Combining dengue control with all Integrated Vector Management program resources is advocated as well as integration of vector control and vaccination strategies.

Before the arrival of chikungunya and Zika, dengue was a serious public health threat. Because such viruses have little clarity in their pathology, many countries without sophisticated public health protocols will not be prepared to manage it. Early detection of mild symptoms of cases in affected areas is crucial to marshal the needed resources to treat infected individual; isolating them to prevent further spread of the disease, to help to decrease the number of severe cases.

CDC has three focus area as a response strategy towards Zika: 1. bring awareness and protection to pregnant women; 2. Control the mosquito vectors, and; 3. expand access to the full range of voluntary contraceptive options for women and men. Methods used include health messaging in media, educational materials distributed at such locations as community and health-care facilities, and weekly surveillance of the disease. Risk-reduction methods include: wearing clothing that covers most of body, using insect repellent adequately, using barriers to keep insects from coming in contact with potential victims, and sleeping under mosquito nets, both day and night.
The location

The metropolitan city of Santiago de Cali, Colombia (hereafter referred to as Cali) covers 560.3 square kilometers and is located at 3°27’26"N and 76°31’42"W. Cali is 1,070 meters above sea level, three hours from the pacific coast of Central Colombia in the department of Valle de Cauca. Cali’s population was 2,181,317 in 2010. The major ethnic groups are Afro-Colombian (26%), Caucasian and mixed (73.3%), and indigenous groups (0.5%). The national annual median gross national income per capita is US$7,560. Just more than twenty-six percent of the population in Cali lives in poverty. The primary cause of death is homicide (including traffic accidents) followed by cerebrovascular events. The median life expectancy in 2010 was 74. The climate is tropical with a median temperature of 24.6°C and annual precipitation of 1,588 mm.

Cali provides an excellent study location to show the importance of a granular perspective on dengue, chikungunya and Zika as it has long been recognized as having a mosquito vectored disease problem. Cali not only mirrors regional trends across the Americas, but locale-specific susceptibility is exacerbated by the violence experienced in Colombia, which disrupted health protection. The city has considerable variation in neighborhood type, and a rich history of dengue-related research. These factors permit researchers to consider the various aspects of disease presence, and the associated risk factors to better understand disease patterns, and therefore develop potential interventions, in such a complex urban environment.

Methods

Incident case data was collected retrospectively from the Colombian Secretary of Health’s national database, SIVIGILA, The National System of Vigilance in Public Health, to which clinics send data on reportable diseases. The MOH combines, cleans, and publishes these data for public use.
Mapping

Different geospatial approaches have been used to look at vector-borne diseases at different scales \(^{68, 69, 70}\) – from using remote sensing and spatial analysis to identify regional / national patterns \(^{71, 72, 73}\), to using more localized spatial analysis \(^{74}\).

Spatial statistics is an area of statics incorporating geographical information. Some methods of spatial statics include cluster analysis \(^{75}\), hotspot analysis \(^{76, 77}\), geographically weighted regressions \(^{72, 78}\), and kernel density analysis \(^{79, 80}\). These methods can be used to statistically analyze and describe epidemiological events with geographical components, and have explicitly interesting applications to vector borne diseases \(^{68, 70}\).

Spatial exploratory analysis of data allows for identification of initial patterns in both surveillance and potential exploratory variables. These exploratory methods can be point or polygon based. For example, a common exploratory method used in vector studies is kernel density analysis (KDA). KDA is a non-parametric method to estimate the probability density function to smooth density data and make population based inferences. KDE is a commonly applied technique in epidemiology. Several examples use mosquito related data, including previous research in Cali, Colombia \(^{66, 81, 82}\). KDE has also been previously used to map out SV and SVG \(^{83, 84}\). Using KDA, surface heat maps can be created to visualize hotspots \(^{85, 86}\). Hotspot analysis is more robust in that case data is first aggregated to a bandwidth and then parametric analyses are run to determine if areas have more or fewer cases than expected by chance.

Hotspots are geographical clusters of cases which exceed the quantity expected to be found at random. Hotspot analysis can be used to identify areas of interest for environmental coding. Hotspots can be identified and statistically confirmed using various methods, for example the Moran’s I test for spatial autocorrelation (similarity of clustered values) \(^{87}\).
The identification of geographic areas of interest may be enough to reveal localized patterns of
disease. However, the next step is to investigate process. One approach involves different forms of
spatial regression, such as Geographically weighted regressions (GWR)\(^\text{88, 89, 90}\). GWRs allow regression
coefficients to vary spatially. Semiparametric models allow some covariates to remain fixed while others
vary spatially.

**GWR Modeling**

Modeling disease incidence can help target interventions to areas\(^\text{91}\). Populations at risk,\(^\text{92}\) and
climate conditions\(^\text{93, 94}\). Models can also determine the reproduction number\(^\text{95}\) and use non-traditional
data to map risk\(^\text{96}\). This is useful for control if models are made free and openly available\(^\text{97, 98}\).

One review of mathematical models of dengue, where 42 studies were systematically reviewed
and critiqued, concluded that “limited understanding of the factors which influence dengue
transmission as well as limited data availability remain important concerns when applying dengue
models to real-world decision problems”\(^\text{99}\). The modelers advocate for more use of vector-host
transmission models which will be useful for projections of combined vaccination and vector control
interventions. In a workshop on dengue modeling hosted by the WHO, it was concluded that
mathematical models are important tools that can be useful to assess the effectiveness of various
intervention strategies\(^\text{100}\).

Traditional models of epidemics are based on susceptible, infected, and recovered individuals.
Other parameters of interest for dengue and vector-borne diseases modeling include serotypes,
immunity, vectors, vaccinations and interventions\(^\text{100}\), human and vector interaction\(^\text{101, 102, 103}\), vector
density\(^\text{104}\), and human mobility\(^\text{105}\). Previous studies have estimated risk by neighborhood in Cali\(^\text{91}\) but
are missing the longitudinal aspects, climate data, and comparisons to Zika and chikungunya. Here,
statistical models of dengue, Zika, and chikungunya incidence are presented as predicted by space, time,
and demographic variables at the neighborhood level. These estimates are useful at the policy level to
allocate limited resources and make decisions based on expected number of cases but these cannot predict outcomes at the sub-neighborhood level nor do these describe the transmission dynamics.

We used an iterative process to build our final GWR models. At each stage, predicted versus residual plots and correlation matrices were inspected to check for serial autocorrelation. Outliers were inspected, transformations made, and covariates added or removed to develop stable models before moving to the next stage.

*Spatial Video Geonarratives (SVG)*

Dengue has generated a large spatial research body, mostly focused on various components of the spatial and spatial temporal patterns of either mosquito intensity or human case data \(^{106, 107}\). Here the lesser covered topics of sub-neighborhood granularity and mixed-method contextual insight are the focus. The need for a more granular understanding at a sub-neighborhood scale that incorporates specific features such as houses, streets, standing water and other disease promoting situations has been identified and was previously acknowledged as an important next direction \(^{64, 108}\). The addition of human context or insight is understood to be important as it has been identified in other spatial research areas \(^{81, 109, 110, 111}\). The combination of these considerations can help explain why spatial variation occurs between neighborhoods of an endemic city. Here, two new epidemiological techniques are further developed, which can be used to address these topics, and in so doing, enrich an already vibrant research agenda focused on dengue in Cali, Colombia.

A core component of most spatial research of mosquito vectored disease involves the identification and explanation of geographic patterns of mosquitoes, and especially disease carrying mosquitoes, or human surveillance data \(^{62, 111, 112, 113, 114}\). Though there are accepted universal risk factors for dengue, such as human density, various measures of economic hardship \(^{62, 81}\), and proximity to urban environmental “risk” features such as canals and ditches, these vary in importance across the same urban space due to variations in the “attributes” of each. An example of such is how the micro-
environments of the same canal will change risks in proximity. Intervention methods such as mosquito control and disease education strategies add spatial and temporal dynamism to these patterns.\textsuperscript{115, 116, 117} Spatial research at a more granular scale of intervention, especially in combination with on-the-ground perceptions and insights can help reveal nuances that will lead to the more general map of risk.\textsuperscript{118} This granular scale is the focus in showing how street level environmental data extracted from spatial video (SV) and Google Street View (GSV), and mapped expert insight captured by spatial video geonarratives (SVG) can be used at any location to better understand disease risk.

Spatial and spatio-temporal patterns of disease, environmental and human correlates and causations, and access to health care and mobility, have all be considered with regards to explaining Dengue in Cali. For example, Delmelle and colleagues\textsuperscript{91} used geographically weighted regression to tease out neighborhood-scale variations within the city. They acknowledge that richer local and contextualized data are required and there are scale issues to be considered as their work was forced to aggregate to appropriate census units. They assert that there are likely to be surveillance deficiencies based on some of those variables found to elsewhere to be predictors of vulnerability. This is of particular concern for urban environments like Cali if dengue vulnerable neighborhoods also coincide with perceived and actual security risks\textsuperscript{59, 119} which may limit the work of vector control in the most at-risk neighborhoods. Additionally, human case data in Cali may be skewed against economically deprived areas where clinical testing may be less frequently used due to a lack of health care access or a lack of confidence in the health care system. While techniques such as Geographically Weighted Regression have been used to gain insight into such neighborhood heterogeneity\textsuperscript{81}, such are still limited by the coarse nature of data input because data have to be standardized, most frequently to a census enumeration unit. Yet all canals, all parks, and neighborhoods of poverty will vary geographically in terms of localized risk. GWR is a useful tool to provide a first analysis, but should be considered as only a first step in identifying sub-neighborhood spaces for more detailed investigation.
Of equal importance is the community perspective which contextualizes risk. The SVG method captures local environmental data and “expert” insight during data collection. Software written at Kent State can be used to produce mapped layers from these transcribed commentaries. If there are enough SVGs, key words can be queried out as a point layer and used as input for more traditional spatial analysis. Alternatively, a single expert could cover multiple spaces and point to institutional insight where none exists. In this case, these narratives can be used to gain further granular insight into a neighborhood or area known to be a hotspot, with the identification of specific features or even behavioral / political actions that increase vulnerability.

There are three types of insight gained from SVG: **Spatially specific**, where an exact location is identified, such as a collection of abandoned tires; **Spatially proximate**, a general risk space identified from a location's sights, sounds, and smells, even though no precise place is identified; and **Spatially inspired or fuzzy spaces**, which contain information that emerges during the SVG ride about the topic under investigation without referencing spatially specific locations. Even with these, there can still be a geographic stimulus – for example, being in an economically disadvantaged neighborhood could result in comments on vector control or disease surveillance, *for that type of neighborhood.*

**Ethics**

This study was approved by the Research Ethics Committee of the participating institutions: Kent State University (#15-529) and University ICESI (#061) (see appendices). All participants completed informed consent and agreed to audio and video recordings. Community health workers were paid 40,000COP ~ $13 USD per interview.
Significance and Statement of goals

As previously described, between September 2014 and April 2016, there were outbreaks of chikungunya and Zika in Cali resulting in continued transmission. Transmission occurs simultaneously with dengue by the same vector in the same target populations.

Our research hypothesis was that dengue, chikungunya, and Zika cases and risk factors have a spatial component. Our goal was to predict future disease patterns and create maps that could be used to direct future control efforts. Three specific aims were the foci: 1. To identify spatial hotspots of dengue, chikungunya, and Zika in Cali Colombia from October 2014 through April 2016 and describe the epidemiology of those cases using G-statistics and classical epidemiological measures; 2. to model the specific risk factors at the neighborhood level to predict dengue, chikungunya, and Zika risk using global, longitudinal, and geographically weighted regression; and 3. to conduct Spatial Video Geonarratives in those hotspots identified in aim 1 to create granular neighborhood level information about environmental and social risk factors for dengue, chikungunya, and Zika transmission.

Impact

The management of vector-borne diseases is costly and often focused on areas of high risk. The maps produced here could be used to identify and control foci of dengue, chikungunya, and Zika in the city. Areas most in need of the type of insights to be gained here also tend to be data poor, and resource (technology, skillsets) limited. In this study, most of the hotspots of interest occurred in areas classified as social strata 1 or 2 on a scale of 1 to 6 with 1 being the most resource poor. The social strata system, as defined by the Congreso de Colombia, classifies homes and neighborhoods by socioeconomic status and is designed to subsidize social services such as utilities and education in low-income areas. Unfortunately, these same social strata often predict quality of life in terms of health, education, and crime with lower social strata corresponding to lower quality of life.
Together with this, a novel method of studying the epidemiological and geographical risk factors associated with dengue and chikungunya in Cali, Colombia was carried out. This study design made use of local expertise, environmental risks, and case data to describe and prescribe methods to prevent future outbreaks, while answering our research questions. Due to the high-risk nature of the endemic neighborhoods, these risk factors could not otherwise be captured so efficiently. Additionally, SVGs provide contextual insight into more traditional hotspot spatial analysis and even provide new data including local expertise and context.
Chapter 2:

Aim 1

Synopsis

Epidemiology of dengue, chikungunya, and Zika in Cali, Colombia during outbreaks of chikungunya and Zika with ongoing dengue transmission. We find that the burden varies by sex and region with maps of incidence presented by neighborhood and incidence rates presented by demographic strata and incidence ratios presented by sex.

Manuscript I:

Epidemiology of chikungunya, dengue, and Zika in Cali, Colombia October 2014-April 2016.

Abstract (250 words)

Here the epidemiology of Zika, chikungunya, and dengue for the period October 2014 through April 2016 in the urban area of Santiago de Cali, Colombia is described. We found that the difference between male and female dengue incidence diminished with the introduction of chikungunya and Zika into the population. Dengue incidence was higher than expected in 3-10 of 2016. This increase may be due to increased reporting due to the Zika outbreak. Severe dengue and dengue mortality represented only 0.8% of all dengue cases. 25% of reported dengue cases were lab confirmed. Chikungunya incidence peaked in 2015 and decreased in 2016. Zika incidence peaked in weeks 9 and 10 of 2016. We see the crude spatial patterns in incidence by disease and neighborhood emerging, indicating that more sophisticated geographically weighted regression may be necessary to predict cases. These results help us to better understand the epidemiological relationship between these diseases spread by a common vector.
Background

Zika \(^1\), dengue\(^2,3,4,5\) and chikungunya\(^6,7\) are vector borne diseases transmitted by bites of an infected *Aedes aegypti* mosquito. Zika can also be transmitted through sexual contact\(^8,9,10\). Individually, each represents an important, global public health threat.

Chikungunya

An estimate of 1.3 billion persons live in areas where chikungunya can be transmitted\(^11\). Chikungunya was responsible for 1,386 - 1,081,962 years of life lost in 2005\(^12\). Chikungunya fever is caused by an alphavirus and characterized by fever and severe joint pain. Chronic arthritis often results but has not yet been consistently defined\(^13,14,15\). There is increasing global concern over expanding distribution\(^16,17,18,19,20,21\). Chikungunya entered the Americas through the Caribbean and quickly spread throughout the Americas\(^122\). Since introduction in 2013, CDC estimates more than 1.7 million suspected cases in 45 countries in the Americas\(^19\). It should be noted that there is confusion between probable and confirmed cases\(^22\) as the disease is often co-endemic and confused with dengue fever\(^18,23,24\) and Zika\(^25\).

Dengue

Dengue reemerged in the Americas in the 1960’s, spread very quickly, became a global threat. Dengue is currently endemic in many regions due to widespread *Aedes* distribution after WW2. The disease is responsible for 390 million infections per year\(^26\). Dengue fever is caused by one of four dengue virus (dengue) serotypes from the family *Flaviviridae* and is characterized by fever, rash, headache, and body aches and in severe cases, shock and hemmorhage. The *Aedes aegypti* mosquito is known to be the main transmitter of the virus, with *Aedes albopictus* as another known vector, with an estimated of 100 million cases a year worldwide\(^28\). Dengue severity is classified as without warning signs, with warning signs, and severe\(^27\).
Women report dengue more often, perhaps due to reporting bias. Severe dengue and mortality from dengue is more frequent in the pediatric population. Elderly patients present complications and more severe dengue.

Zika

Zika virus entered the scene when mosquitoes distribution and human travel were prime factors in creating an epidemic that spread very quickly across the globe. Zika is a reemerging vector-borne viral disease declared to be a public health emergency of international concern in 2016 by WHO director general, Margaret Chan, which was reclassified as a continued threat November 18, 2016. Zika virus is a flavivirus that causes acute febrile illness and is associated with microcephaly and other neurological disorders. It is transmitted by the same vectors as dengue and chikungunya, Aedes albopictus and Ae. aegypti. The symptoms it presents are minor and last a maximum of a week, with some cases not showing any symptoms at all. 25% of infected individuals exhibit symptoms lasting 2-7 days, which include mild fever, rash, conjunctivitis, and muscle pain. In epidemic areas, an increased number of neurological syndromes have been observed including Guillain-Barré Syndrome (Guillain–Barré Syndrome), meningoencephalitis, and myelitis. Zika has also been associated with congenital microcephaly and other birth defects with increased incidence of microcephaly observed in children born to pregnant women infected with Zika. According to WHO, the virus is known to circulate in Africa, the Americas, Asia and the Pacific with recent outbreaks in the Americas.

Women report more Zika even after adjusting for routine testing of pregnant women, presumptively due to sexual transmission from men to women.

Control

Despite recent increases in reported outbreaks and distribution, no vaccines or cure for these diseases exist. New vaccines, however, are under development. The Zika virus vaccines are showing
promise with plans for clinical trials. Some nucleoside analogue drugs have shown promising cellular activity. Chikungunya has no licensed vaccines but approximately 15 candidates are in the process of being developed and several others are undergoing clinical trials. The first dengue vaccine was licensed for those age 9-45 years of age living in endemic areas, and two more are beginning phase three of investigation in 2016. No specific treatment exists for any of these diseases, rather the focus is on relieving the patient’s symptoms.

Vector control for the Aedes mosquito vector remains the primary prevention method. Common risk factors for dengue, chikungunya, and Zika are environmental as the Aedes mosquito, can breed in any container of stagnant water (e.g. flower pots, puddles in road, tree holes, tires, bromeliads, buckets, drums, small trash, bottles, cans) and are well adapted to urban environments. Traditional methods of vector control include indoor residual sprays, fumigation, and source reduction. Biological forms of vector control are gaining popularity. For example, Sterile Insect Techniques releases sterilized mosquitoes, to reduce the population through failed reproductive events. In Medellin, Colombia, Wolbachia infected Ae. aegypti are being released and have reduced vector competence for ZIKV.

Combining dengue control with all Integrated Vector Management program resources is advocated as well as integration of vector control and vaccination strategies studies. Before the arrival of chikungunya and Zika, dengue was a serious public health threat. Countries not prepared to have such viruses with little clarity in their pathology, will not be prepared to manage it. Early detection of mild symptoms of cases in affected areas is crucial to marshal the needed resources to treat individuals and isolating them to prevent further spread of the disease, which would decrease severe cases.

CDC has three focus areas as a response strategy towards Zika: 1. Bring awareness and protection to pregnant women; 2. Control the mosquito vectors; and 3. Expand access to the full range of voluntary contraceptive options for women and men. Health messaging in media, education material
has been implemented have been distributed at locations such as communities and health care facilities, and weekly surveillance of the disease. Risk reduction methods include wearing clothing that covers most of the body; using insect repellent adequately; using barriers to keep insect from coming in contact; and sleeping under mosquito nets, both day and night.

Reducing the vector population density through focused, customized, and integrated vector control can reduce vector-borne disease transmission including that of dengue, chikungunya and other emerging vector-borne diseases in the region such as Zika and promote healthy environments. Various vector control methods exist and application should be customized based on local vector and logistics. In general, the most effective methods reduce the lifespan of adult mosquitoes by exponentially reducing vectorial capacity, the number of infective bites received daily by a single host. One review of the epidemiology in Colombia concluded that “there has been no success in the effective control of the disease.”

Estimates from Brazil found that the abundance of mosquitos is lower and more clustered during winter compared to high abundance periods where distribution is more homogenous.

The location

The metropolitan city of Santiago de Cali, Colombia (hereafter referred to as Cali) covers 560.3 Km², is located at 3°27'26"N and 76°31'42"W, is 1,070 meters above sea level, 3 hours from the pacific coast of Central Colombia in the department of Valle de Cauca with a population of 2,181,317 in 2010. The major ethnic groups are Afro-Colombian (26%), Caucasian and mixed (73.3%), and indigenous groups (0.5%). The national annual median GNI per capita is US$7,560 with 26.1% of the population in Cali living in poverty. The primary cause of death is homicide (including traffic accidents) followed by cerebrovascular events. The median life expectancy in 2010 was 74. The climate is tropical with a median temperature of 24.6°C and annual precipitation of 1,588mm.
During the study period, the city was also experiencing El Niño weather, severe droughts, flooding, a new mayor, a budget shutdown of the local public hospital, the promise of a new dengue vaccine soon to be available, and two new diseases: chikungunya and Zika.

Chikungunya in Cali

Colombia experienced an outbreak of chikungunya between September 2014 and September 25, 2015. Per the Colombian Ministry of Health and Social Protection, 439,000 cases accumulated in 712 municipalities during the outbreak.

Cali was one of the most affected regions in the country with 44,877 accumulated cases reported up to October 17, 2015 due to the Asian genotype. As to our knowledge, no detailed map of this outbreak exists to date. 63.6% of chikungunya cases were reported among females and 11.3% among 25-29-year-olds.

The 2014 outbreak of chikungunya cost an estimated US$73.6 million due to 40.44 to 45.14 lost/100 000 population DALYs lost, highlighting the effects of acute and chronic impacts if left uncontrolled. The 2014-2015 chikungunya outbreak affected 1.5 million people in the Americas with cases of atypical mortality reported in Colombia.

The main symptoms of the chikungunya outbreak in Colombia were arthralgia, fever, and rash with the chikungunya sequences detected within the Asian genotype and closely related to the British Virgin Islands strain with some fatal cases reported. Chikungunya vertical transmission rate ranged between 27.7% and 48.29%. The case fatality rate (CFR) at the only center that reported deaths was 5.3% (Asian genotype of chikungunya).

A meta-analysis of CHIK cases estimates that 25-34% will develop pCHIK-CIR are and 14% chronic arthritis. In one prospective study from Colombia, neonates are more susceptible to severe chikungunya and present higher viral loads. Post-chikungunya polyarthralgia presents in 44.3% of chikungunya cases on average. Estimates in Colombia for post-chikungunya chronic inflammatory...
rheumatism were 30.61-34.04 DALYs per 100,000 in 2014 compared to 25.45-28.31 in all Latin America.

DALYS: Post-chikungunya chronic inflammatory rheumatism (pCHIK-CIR) after 2014 outbreak which will be 2/3 the burden due to heart disease in Colombia. Post-chikungunya chronic inflammatory rheumatism (pCHIK-CIR) frequency ranges from 14.4% to 87.2%. Model estimated a prevalence of 47.57% for pCHIK-CIR (95% CI 45.08–50.13), with a median time to 50% of pCHIK-CIR in 20.12 months. Local clinics must deal with these cases after the acute phase. Half of patients diagnosed with Post-chikungunya develop chronic inflammatory rheumatism.

Congenital chikungunya virus (CHIK) infection is infrequent but can result in infant mortality as reported in one case series from Colombia (3/7) with the clinical and epidemiological features of congenital and neonatal cases of CHIK not well defined in LAC.

_Dengue in Cali_

Dengue is hypoendemic in Colombia and Cali with over 18,000 cases reported to SIVIGILA, The National System of Vigilance in Public Health, in the department of Valle de Cauca between week 1 and week 48, 2015.

In the year 2010, there was a major dengue outbreak in Colombia, with more than 150,000 cases and 289 deaths reported in the country. More than one million cases were reported in all Latin America. 50.9% of dengue cases were recorded in females with 58.5% being reported from the public insurance contributory scheme and 2.7% from persons without affiliation. 24.8% of dengue cases were seen in persons <= 15 years of age, while 34.1% presented as severe dengue. It has been noted that 725 cases were seen in indigenous groups and 1,643 cases in Afro-Colombians. In Cali, Colombia, a week in July, 2016 recorded a total of 126 cases, which is an improvement from the situation the city was facing in the last months, wherein the number reached 31,001.

A systematic review of epidemiology of dengue in Colombia concluded that “dengue disease in Colombia was characterized by a stable “baseline” annual number of dengue fever cases, with major
outbreaks in 2001–2003 and 2010. The geographical spread of dengue disease cases showed a steady increase, with most of the country affected by the 2010 outbreak.” 137. They also reported that case fatality rates were high in the 1990s, dipped from 2000-2009, and increased greatly in 2011. The geographical range increased from 2000-2010. Most dengue occurs in those <15 years of age, with a majority of cases in children <1 in 2009 162. In the department of Valle del Cauca, the number of cases has been increasing from 2010 to 2015, 163. possibly due to improvements in diagnosis and clinical management 162.

Dengue and chikungunya differential diagnosis in pediatric patients is not reliable and requires lab testing 164. Health care professionals should be updated on symptoms and epidemiology of chikungunya 165. There are new, easy, and inexpensive methods of molecular detection of dengue, chikungunya and Zika viruses for febrile patients, which are especially in areas of co-circulation 166. Porrino et al argue that Zika, and similarly other neglected diseases, may have been circulating undetected in Colombia prior to the outbreak and simply misdiagnosed 167.

*Zika in Cali*

Zika was imported to Colombia from Brazil with the first new cases reported in October, 2015 168 and 20,297 suspected and 1,050 laboratory confirmed cases reported by the National Institute of Health between epidemiological week 40 of 2015 and week 3 of 2016 169. Through April 2016, 3,139 suspected cases of Zika were reported to the Secretary of Health in Cali 67. 66.10% of reported Zika cases were females with 63.57% from the contributory scheme and 14.09% among 25 to 29-year-olds, 2.66% among infants one-year-old or less, and 3.24% among adults 65 or older. 0.68% were reported among indigenous groups and 1.83% among afro-Colombians 144.

The introduction of Zika created the perfect trifecta of infection in Colombia 170, 171 and across Latin America. Despite co-circulation of dengue, Zika, and chikungunya in Latin American countries172, among those patients that were laboratory tested for all three, most are only diagnosed with one. We
are not sure if previous infection with dengue and chikungunya are a risk factor for severe Zika. More data is needed 172.

Reported attack rates reported in San Andres and Girardot, Colombia were 12.3-18.43 per 1,000 residents respectively with higher rates reported in females and 20-49 age group. The R0 was estimated at 1.41 in San Andres and 4.61 in Girardot, Colombia 173.

Colombia declared a Zika outbreak in October, 2015. Surveillance started August 2015. 65,726 cases reported in Colombia by April 2, 2016, (4% confirmed by RT-PCR) with females reporting 2x as much as males. 11,944 pregnant women reported with 14% confirmed by RT-PCR. No abnormalities were reported among infants born to mothers infected in the third trimester (n = 1850, 90% followed to birth). Four microcephaly cases had laboratory evidence of congenital Zika. “Preliminary surveillance data in Colombia suggest that maternal infection with the Zika virus during the third trimester of pregnancy is not linked to structural abnormalities in the fetus.”174.

58,838 cases of Zika were reported in Colombia from September 22, 2015-March 19, 2016, with cases of Zika associated mortality and severe Zika reported. A rapid risk assessment is needed for clinical settings 175. Zika virus associated deaths were reported in Colombia: “we report Zika virus infection in Colombia in association with an ongoing outbreak of acute maculoexanthematous illness. Since detection of Zika virus in Sincelejo, a total of 13,500 cases have been identified in 28 of the country’s 32 territorial entities, all of which have abundant populations of Ae. aegypti mosquitoes and co-circulation of dengue and chikungunya” 176.

There is evidence that Zika is associated with Guillain–Barré Syndrome in Colombia 177,178. Estimated 0.08% probability of Guillain-Barré syndrome in Zika-infected individuals in Colombia 179. Among 796 cases of Guillain–Barré Syndrome, 31% reported infectious disease prior to Guillain–Barré Syndrome onset (most common were gastrointestinal and respiratory. 15% of cases resulted in mortality. However, data on Guillain–Barré Syndrome is not sufficient and more data should be
collected considering the association of Zika and Guillain–Barré Syndrome. Guillain–Barré Syndrome associated with Zika virus infection in Colombia at a rate of 90 cases per month on a national level during the outbreak compared to 20 cases per month 2009-2015. Reports of Guillain–Barré Syndrome during chikungunya outbreak might indicate that Zika was circulating earlier. Increased incidence of Guillain–Barré Syndrome was reported during Zika outbreak (2-9.8 times higher than baseline) in 7 countries including Colombia.

Bar-Yam, Costa, et al found that there were fewer microcephaly cases reported than expected, compared to rates of microcephaly in French Polynesia (6 reported through June 11, 2016 compared to 200 expected total). Brazilian officials have questioned whether Zika acts alone to cause birth defects with other SES factors possibly important.

Control in Cali

Colombia is advising its territories to implement a plan to control breeding sites, note increases in outbreaks in their communities and to analyze fatal cases. Additionally, they urge territories to inform the communities of the symptoms and signs of alarm of dengue and to advise persons to seek medical attention when needed.

The secretary of Valle del Cauca in Colombia exhorts people to understand that they need to live with the mosquito and to fight against it with preventive measures. Also, the ministry advises pregnant woman to take precautions. It was also stated that the local community has a very important part in eradication of the Aedes aegypti mosquito.

The city of Cali has provided guidelines to multiple organizations involved in order to keep the epidemic controlled which include the elimination of breeding sites, surveillance over possible cases and to report these cases immediately. During the data collection period, the authors note an increase in vector control during the Zika outbreak in the affected communities focusing on household fumigation and storm drain cleaning as reported in the local newspaper.
Previous intervention in the study area tried to remove intra-domiciliary breeding sites with no evidence of success, suggesting that outdoor breeding sites may be more important. Screens were a protective factor and presence of water plants and flower pots were risk factors. There is evidence of two lineages of *Ae. aegypti* in Colombia, one present in areas of high dengue incidence and insecticide use and one in low incidence and insecticide areas. Wolbachia infected *Ae. aegypti* are being released in Medellin, Colombia and have reduced vector competence for ZIKV.

Recommendations to delay pregnancy during outbreak were issued but the edict came with the caveat that this method could exacerbate the outbreak if not timed correctly. Also, there are concerns that Hormonal contraceptive prescriptions have not increased in Colombia since the recommendation by the Colombian MOH in January 2016, although this may be related to high rates of unwanted pregnancies and related factors (access to contraception, control over reproductive health, quality of sex education).

**Materials and Methods**

The study design consisted of a mixed methods retrospective and prospective format. Incident case data was collected retrospectively from the secretary of health. The study population included cases of confirmed or suspected chikungunya, dengue or Zika from the municipal area of Cali, Colombia reported between October, 2014 and April, 2016 to the national database, SIVIGILA (The National System of Vigilance in Public Health). Incarcerated individuals were excluded.

**Statistical analysis:**

Statistical analysis was performed using Stata 12. We described each disease class by subgroups of severity, age, sex, ethnicity, lab confirmation, and social risk group. We performed univariate analysis to determine the behavior of the numeric variables and used Shapiro Wilk test of normality of the variables and those with $p > 0.05$ were considered normally distributed and a mean and standard deviation calculated. For non-normal variables, median and interquartile ranges were calculated.
Categorical variables are presented as proportions and strata compared with chi-squared tests with Fisher’s exact test used for tables with values less than five in any cell. Outliers were removed for final analysis.

Strata specific weights were made by year, age, and sex for all incidence calculations using census data. Chikungunya cases were estimated to have been reported for 5% of the total burden at the individual level and were weighted accordingly (MOH personal communication).

Counts of disease over time were mapped by neighborhood and animated over time using the Surveillance and animation packages respectively in the open source statistical package R studio Version 1.0.44. We calculated the cumulative incidence by case age, sex, ethnicity, social risk group, and occupation stratified by disease as a ratio of number of cases diagnosed during the period of interest and the denominator as the population of Cali by strata and incidence was estimated over time by week, month, and quarter. We used weighted incidences to produce incidence maps by neighborhood using EPSG Projection 3115 - MAGNA-SIRGAS / Colombia West zone projection in ArcGIS 10.3.
Results

Demographic

Table 1 shows the sample demographic characteristics by disease and disease severity. 4,636 out of 21,210 (21.86%) reported dengue cases were laboratory tested by indirect IgM, IgM, PCR, or NS1 for dengue of which 1,843 were positive (39.75%). We did not have access to laboratory results for Zika or Chikungunya. There were differences in dengue severity category reported across age, gender, travel within the last 15 days, and pregnant vs non-pregnant females.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Level</th>
<th>Dengue w/o warning</th>
<th>Dengue w/ warning signs</th>
<th>Severe dengue or death</th>
<th>Dengue unclassified</th>
<th>Chikungunya</th>
<th>Zika</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Lab result**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>9963</td>
<td>2405</td>
<td>158</td>
<td>8684</td>
<td>2636</td>
<td>3139</td>
<td></td>
</tr>
<tr>
<td>(10.4)</td>
<td></td>
<td>(31.6)</td>
<td>(29.1)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>939 (9.4)</td>
<td>406 (16.9)</td>
<td>10 (6.3)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Compatible</td>
<td>1 (&lt;1)</td>
<td>2 (0.1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Doubtful</td>
<td>4 (&lt;1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Inadequate</td>
<td>1 (&lt;1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>No Data</td>
<td>6739 (67.6)</td>
<td>1055 (43.9)</td>
<td>97 (61.4)</td>
<td>8683 (100)</td>
<td>2636 (100)</td>
<td>3139 (100)</td>
<td></td>
</tr>
<tr>
<td>Non-Reactive</td>
<td>0 (0.0)</td>
<td>1 (&lt;1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Not Compatible</td>
<td>2 (&lt;1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Not processed</td>
<td>4 (&lt;1)</td>
<td>4 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Undefined</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>2 (1.3)</td>
<td>1 (&lt;1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Value registered</td>
<td>1236 (12.4)</td>
<td>177 (7.4)</td>
<td>3 (1.9)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Age Category **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-14</td>
<td>1956 (19.6)</td>
<td>860 (35.8)</td>
<td>59 (37.3)</td>
<td>2299 (26.5)</td>
<td>300 (11.4)</td>
<td>377 (12.0)</td>
<td></td>
</tr>
<tr>
<td>15-34</td>
<td>4125 (41.4)</td>
<td>823 (34.2)</td>
<td>48 (30.4)</td>
<td>3542 (40.8)</td>
<td>834 (31.6)</td>
<td>1334 (42.5)</td>
<td></td>
</tr>
<tr>
<td>35-64</td>
<td>3264 (32.8)</td>
<td>510 (21.2)</td>
<td>23 (14.6)</td>
<td>2397 (27.6)</td>
<td>822 (31.2)</td>
<td>1294 (41.2)</td>
<td></td>
</tr>
<tr>
<td>65+</td>
<td>618 (6.2)</td>
<td>212 (8.8)</td>
<td>28 (17.7)</td>
<td>446 (5.1)</td>
<td>680 (25.8)</td>
<td>134 (4.3)</td>
<td></td>
</tr>
<tr>
<td>Gender*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>4776 (47.9)</td>
<td>1219 (50.7)</td>
<td>70 (44.3)</td>
<td>4346 (50.0)</td>
<td>1710 (64.9)</td>
<td>1986 (63.3)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5187 (52.1)</td>
<td>1186 (49.3)</td>
<td>88 (55.7)</td>
<td>4338 (50.0)</td>
<td>926 (35.1)</td>
<td>1153 (36.7)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indigenous</td>
<td>29 (0.3)</td>
<td>14 (0.6)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>8 (0.3)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>ROM, Gitano</td>
<td>17 (0.2)</td>
<td>4 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>9 (0.3)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Raizal</td>
<td>11 (0.1)</td>
<td>6 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>5 (0.2)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Afro-Colombian</td>
<td>222 (2.2)</td>
<td>110 (4.6)</td>
<td>3 (1.9)</td>
<td>1 (&lt;1)</td>
<td>61 (2.3)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>9684 (97.2)</td>
<td>2271 (94.4)</td>
<td>87 (55.1)</td>
<td>10 (0.1)</td>
<td>2172 (82.4)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>No data</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>68 (43.0)</td>
<td>8673 (99.9)</td>
<td>381 (14.5)</td>
<td>3139 (100.0)</td>
<td></td>
</tr>
<tr>
<td>Disability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>8715 (87.5)</td>
<td>1966 (81.7)</td>
<td>81 (51.3)</td>
<td>11 (0.1)</td>
<td>2071 (78.6)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5 (0.1)</td>
<td>4 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1 (&lt;1)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>No data</td>
<td>1243 (12.5)</td>
<td>435 (18.1)</td>
<td>77 (48.7)</td>
<td>8673 (99.9)</td>
<td>564 (21.4)</td>
<td>3139 (100.0)</td>
<td></td>
</tr>
<tr>
<td>Displace</td>
<td>No</td>
<td>8720 (81.8)</td>
<td>81 (51.3)</td>
<td>11 (0.1)</td>
<td>2065 (80.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes (%)</td>
<td>No (%)</td>
<td>No data (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------</td>
<td>---------</td>
<td>--------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency in the last 15 days*</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Migrant</td>
<td>77 (48.7)</td>
<td>8673 (99.9)</td>
<td>567 (21.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>433 (18.0)</td>
<td>87.5</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnant females*</td>
<td>458 (5.3)</td>
<td>3812 (43.9)</td>
<td>1287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>983 (40.9)</td>
<td>124 (78.5)</td>
<td>1615</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orphan</td>
<td>81 (51.3)</td>
<td>8673 (99.9)</td>
<td>564</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1972 (82.0)</td>
<td>78.5</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community mother</td>
<td>34 (0.3)</td>
<td>8673 (99.9)</td>
<td>398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>34 (0.3)</td>
<td>8673 (99.9)</td>
<td>398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>26 (1.1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex-para-military/guerilla</td>
<td>11 (0.1)</td>
<td>8673 (99.9)</td>
<td>564</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>33 (1.4)</td>
<td>8673 (99.9)</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>324 (14.8)</td>
<td>8673 (99.9)</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under psychiatric care</td>
<td>367 (15.1)</td>
<td>8673 (99.9)</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1969 (81.9)</td>
<td>87.5</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>326 (14.8)</td>
<td>8673 (99.9)</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Another group</td>
<td>18 (0.8)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Violence victims</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>11 (0.5)</td>
<td>8673 (99.9)</td>
<td>3139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3 (0.2)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Frequency (%), **p<0.001, *p<0.05 for comparison between dengue categories
Cases over time

Dengue peaked in weeks 18, 2015 and 8, 2016 (Figure 1) but chikungunya wasn’t reported until week 1 of 2015, peaking in week 20 of 2015. Reporting of Zika cases started in week 47 of 2015 and peaked in week 10 of 2016. This is the expected pattern with greater incidence in the first and second quarters of the year with the rainy season from November – May followed by a hotter drier period from June – September, although 2015 was an El Niño year causing unusually hot and dry weather (Figure 2).

Figure 1: Weekly and quarterly incidence
For each disease outcome over time, the central western region reported the highest cumulative number cases (ranging from 0-576 for dengue, 0-49 for chikungunya, and 0-81 for Zika). Statistics show, however, that in January 2016, nearly all the city was reporting cases of Zika (http://capturadedatos.co/surveillance/). For dengue, the entire city is endemic over the complete study period, but only the same central western region seems to be disproportionately affected by the peaks of the outbreaks (May 2015 and February 2016). Chikungunya conversely affected all of the city and only peaked in February 2015, again disproportionately affecting the central western region (Figure 3 Click here http://capturadedatos.co/surveillance/# to view the evolution of these diseases in Cali, Colombia.) and use the login id: dengue, Password: losmapasdecali.
Figure 3: a. dengue October 2014 - April, 2016, b. chikungunya January 2015 - April, 2016. C. Zika November 2015 - April, 2016. Click here http://capturadedatos.co/surveillance/# to view the evolution of these diseases in Cali, Colombia. login id: dengue, Password: losmapasdecali

Per 100,000 population, incidence ranged from 0-304 for Zika, 0-2,782 for dengue, and 0-260 for chikungunya (Figure 4). High incidence areas included Ciudad Universitaria, Planta de Tratamiento, Portrero Grande, Acuaducto de San Antonio, and Menga (Figure 4).
Females reported disproportionately higher numbers of cases of chikungunya compared to males in all years (1.4-1.6x) and Zika in 2016 (1.4x) after the connection with microcephaly had been made common knowledge by media and the government of Colombia issued a warning to women to delay pregnancy. Males always reported more cases of dengue compared to women but this difference seems to diminish in 2016 when females presumably starting testing for dengue and Zika at the same time. Especially females of child bearing age reported more chikungunya and Zika compared to males and this affect diminished or disappeared in other groups with males at higher risk (Figure 5).
Figure 6: Distribution of days from onset of symptom to exam by test result.

There also appears to be a delay in reporting from onset of symptoms (median = 6 days) (Figure 6). The ratio of negative to positive was higher on days 0, 2-4, 6 – 8, 20, and 20+. After day 7 the difference diminished or the ratio reversed suggesting that those motivated to take an exam after 6 days of symptoms may be severe or that diagnostics are not catching early disease.

Here the epidemiology of dengue, chikungunya, and Zika incidence in Cali, Colombia from October 2015-April 2016 is described. During this time, the city experienced two outbreaks, drought, rainy seasons, and change at the political level (mayoral elections). We see changes in reporting over time as Zika was introduced and females of reproductive age submitted to test more. This eliminated
the difference previously seen between males and females in dengue incidence, suggesting that the historical difference was due to reporting bias rather than difference in risk. The most at risk groups for dengue were women of reproductive age. The most at risk group for severe dengue were males older than 15. The highest risk seasons were May and February during rainy seasons. This data can help us to target vector control methods in space and time and better support ongoing work in the community to describe the environmental, geographical, and social risk factors.

**Conclusions**

Limitations of this study include the use of secondary passive surveillance data to estimate disease burden. We did not have access to laboratory confirmation for disease state. The sample was biased due to non-random underreporting, perhaps related to poverty and education.

Here the difference between male and female dengue incidence diminished with the introduction of chikungunya and Zika into the population. Dengue incidence was higher than expected in weeks 3-10 of 2016. This increase may be due to increased reporting due to the Zika outbreak. Severe dengue and dengue mortality represented only 0.8% of all dengue cases. 25% of reported dengue cases were lab confirmed. Chikungunya incidence peaked in 2015 and decreased in 2016. Zika incidence peaked in weeks 9 and 10 of 2016. Finally, the crude spatial patterns in incidence by disease and neighborhood emerge, indicating that more sophisticated geographically weighted regression may be necessary to predict cases. These results help us to better understand the epidemiological relationship between these diseases spread by a common vector.
Chapter 3

Aim 2

Synopsis

Manuscript II:

Modeling and predicting chikungunya, dengue, and Zika in Cali, Colombia October 2014-April 2016.

Abstract

Here the counts of disease by neighborhood are modeled using OLS, geographically weighted Poisson regressions, and longitudinal generalized estimating equations using a Poisson model and a log link. All models used percentage male population, percentage afro-Colombian, service anomalies index, average rain and lagged rain, lagged temperature anomalies, social strata, distance to the nearest canal, total population, and surface area by neighborhood. Independent variables were created from the 2005 census, canal shapefiles, satellite rain data, and weather station temperature data. Case counts were secondary case data from SIVIGILA. We compared all three models. No evidence of spatial or temporal autocorrelation. Statistically significant risk factors include: male population percentage, afro-Colombian percentage, social strata, rain, and temperature. These models can be used to better understand the outbreaks that occurred 2014-2016 in Cali, Colombia.

Introduction

Chikungunya, dengue, and Zika are vector-borne diseases important to the burden of public health worldwide and in Colombia. These three present similar fever-like illness and
are transmitted by Aedes species with expanding distribution of infected vectors in Colombia \textsuperscript{197}. There is currently no cure or vaccine despite recent increases in reported outbreaks and distribution.

Cali experienced an outbreak of chikungunya between September 2014 and September 25, 2015 \textsuperscript{139,140}. As to our knowledge, no detailed map of this outbreak exists to date \textsuperscript{142}. Dengue is endemic to Colombia and Cali \textsuperscript{26,159}. Zika was imported to Colombia from Brazil with the first new cases reported in October, 2015 \textsuperscript{168,169}.

Dengue and chikungunya share many common risk factors but dengue has been endemic in Cali since reintroduction to Colombia in 1972 \textsuperscript{198,199} whereas chikungunya was introduced in September, 2014 and exhibited an outbreak pattern. Here these two epidemiological patterns are compared and socio-economic risk factors are investigated to predict disease incidence. We can also compare models excluding and including geographic patterns to determine their impact on the cofactors.

Modeling disease incidence can help target interventions to areas \textsuperscript{91} and populations at risk \textsuperscript{92} and climate conditions \textsuperscript{93,94}. They can also determine the reproduction number \textsuperscript{95} and use non-traditional data to map risk \textsuperscript{96}. This is useful for control if models are made free and openly available \textsuperscript{97,98}.

One review of mathematical models of dengue, where 42 studies were systematically reviewed and critiqued, concluded that “limited understanding of the factors which influence dengue transmission as well as limited data availability remain important concerns when applying dengue models to real-world decision problems”\textsuperscript{99}. However, reviewers advocate for more use of vector-host transmission models which will be useful for projections of combined vaccination and vector control interventions. In a workshop on dengue modeling hosted by the WHO, they concluded that mathematical models are important tools that can be useful to assess the effectiveness of various intervention strategies \textsuperscript{100}. 


Traditional models of epidemics are based on susceptible, infected, and recovered individuals. Other parameters of interest for dengue and vector-borne diseases modeling include serotypes, immunity, vectors, vaccinations, and interventions \(^{100}\), human and vector interaction \(^{101, 102}\), vector density \(^{104}\), and human mobility \(^{105}\). Previous studies have estimated risk by neighborhood in Cali \(^{91}\) but are missing the longitudinal aspects, climate data, and comparisons to Zika and chikungunya. Here statistical models of dengue, Zika, and chikungunya incidence as predicted by space, time, and demographic variables at the neighborhood level are presented. These estimates are useful at the policy level to allocate limited resources and make decisions based on expected number of cases. These cannot, however, predict outcomes at the sub-neighborhood level nor do these estimates describe the transmission dynamics.

**Methods**

This study was carried out using secondary data from Santiago de Cali, Colombia (hereafter referred to as Cali), previously described in the background, from SIVIGILA, The National System of Vigilance in Public Health \(^{67}\). Remotely sensed Daily precipitation at 4 km spatial resolution \(^{200}\) and Monthly weather station temperature anomalies data collected from Alfonso Bonilla Aragón International Airport in the north of Cali \(^{201}\) were used at the monthly scale (Figure 7). Rain data were averaged by month and by neighborhood and added to the model lagged by one month and non-lagged. Temperature added to the model lagged by one month and non-lagged for the longitudinal models. We created indices (Table 2) from census 2005 data \(^{60}\).

Box plots for each covariate were created and service index coverage, proportion single, and proportion male were replaced with the mean if values were extreme outliers.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Level</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td></td>
<td>6113</td>
</tr>
<tr>
<td>Dengue count, median (IQR)</td>
<td></td>
<td>0 (0, 0)</td>
</tr>
<tr>
<td>Zika count, median (IQR)</td>
<td></td>
<td>0 (0, 0)</td>
</tr>
<tr>
<td>Chikungunya count, median (IQR)</td>
<td></td>
<td>3 (2, 5)</td>
</tr>
<tr>
<td>Precipitation (mm), median (IQR)</td>
<td></td>
<td>0.518 (0.474, 0.568)</td>
</tr>
<tr>
<td>Temperature anomaly median (°C), median (IQR)</td>
<td></td>
<td>457 (439, 467)</td>
</tr>
<tr>
<td>Service coverage index, median (IQR)</td>
<td></td>
<td>4957 (2984, 8930)</td>
</tr>
<tr>
<td>Services index, median (IQR)</td>
<td></td>
<td>28.5% (4.2%)</td>
</tr>
<tr>
<td>Educational assistance (%), mean (SD)</td>
<td></td>
<td>6.3% (2.0%)</td>
</tr>
<tr>
<td>Disability (%), mean (SD)</td>
<td></td>
<td>89.68% (86.0, 92.0)</td>
</tr>
<tr>
<td>Literacy (%), median (IQR)</td>
<td></td>
<td>4149 (2388, 7327)</td>
</tr>
<tr>
<td>Educational index, median (IQR)</td>
<td></td>
<td>47120.0 (3020.5, 7916.3)</td>
</tr>
<tr>
<td>Educational assistance index, median (IQR)</td>
<td></td>
<td>1.8% (1.2, 2.9)</td>
</tr>
<tr>
<td>Empty homes (%), median (IQR)</td>
<td></td>
<td>779 (12.7%)</td>
</tr>
</tbody>
</table>

### Table 2: Census 2005 characteristics by neighborhood

**Social Strata**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1648 (27.0%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2299 (37.6%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>494 (8.1%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>684 (11.2%)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>209 (3.4%)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>47.2 (45.6, 48.0)</td>
</tr>
</tbody>
</table>

| Male (%), median (IQR) | 19.0% (11.3, 31.9) |
| Mulato, Afro-Colombian (%), median (IQR) | 0.3% (0.2, 1.1) |
| Unemployment (%), median (IQR) | 14.4% (11.7, 15.8) |
| House keeper, median (IQR) | 38.6% (36.6, 40.5) |
| Single (%), median (IQR) | 100 (99, 100) |
| Houses with plumbing (%), median (IQR) | 100 (99, 100) |
| Houses with energy (%), median (IQR) | 2406.4 (1740.4, 3114.0) |
| Neighborhood area, median (IQR) | 0.52 (0.47, 0.57) |

**Statistical analysis:**

We used an iterative process to build our final models. At each state, predicted versus residual plots and correlation matrices were inspected to check for serial autocorrelation. Outliers were inspected, transformations made, and covariates added or removed to develop stable models before moving to the next stage.
A global Poisson model was estimated in Stata 12 using stepwise regression with entry at $p = 0.1$ and exit at $= 0.2$. Variables found significant in any of the three disease outcome models were considered for inclusion in the final global Poisson models and subsequent models.

A longitudinal Poisson regression to predict number of cases by neighborhood by month was estimated using Stata 12 XT generalized estimating equations using a log link and autoregressive 1 correlation matrix using the Huber sandwich method to estimate standard errors more robust to violations of the linear relationship between log outcome and explanatory variables. These results were compared to a mixed effects model using Stata 12 xtmepoisson model with an independent covariance structure allowing the anomalies of rain, 1-month lagged rain, temperature and 1-month lagged temperature to vary but no difference was found between the two models ($\text{chi}^2(01) = 1$) so the simpler xtmepoisson is presented here.

Incidence of dengue, chikungunya, and Zika were predicted as a count using a semiparametric Geographically weighted Poisson regression with an adaptive bi-square geographic kernel, golden section search for optimal bandwidth, AICc Criterion for optimal bandwidth, with 19 varying coefficients and 0 fixed coefficients. Resulting GWR model coefficients (Figure 9) were visualized using ArcGIS 10.3.1. Neighborhood risks (Figure 8) were described by census 2005 data (Table 2) using STATA 12 and visualized using ArcGIS 10.3.1.

**Results**

We ran a series of models with each disease outcome. First, a global stepwise AIC Poisson regression was run with robust standard errors (1. SWAIC Poisson) and a stepwise Poisson model with an exiting p-value of 0.1 and an entering value of 0.05 with robust standard errors (2. stepwise p Poisson). Those variables that remained in either model were considered for inclusion in subsequent models. Then a global negative binomial regression was estimated with robust standard errors (3. nb global). Then a longitudinal general estimating equation was modeled with a negative binomial
distribution and a log link and a second order autoregressive correlation matrix and robust standard errors (4. xt nb). Finally, a longitudinal negative binomial regression with random intercepts by neighborhoods (5. xt nb random) was fit. The Poisson models displayed overdispersion (>1) giving unreliable estimates so the models are interpreted beginning with the 3rd model, the negative binomial regression.

The global Poisson models (Table 3 models 1-3) are estimated and interpreted here but ignore the cyclical nature of vector-borne disease transmission, especially due to the effects of temperatures and rain on vector habitats and human water storage habits. Of greater importance, during the study period, there were outbreaks of Zika and chikungunya in addition to the hypoendemic patterns of dengue expected (Figure 7).
The estimates shown here are exponentiated coefficients interpreted as incidence risk ratios or the increased risk of outcome compared to the baseline or comparison group.

For Dengue, for each percentage point increase in afro-Colombian population in a neighborhood, the risk ratio increases by 11.7 relative to baseline. This relationship holds true in the longitudinal NB model (4) with an IRR of 18.4 and in the random intercepts model with an IRR of 1.5. This sharp decline in effect suggests that a geographically weighted regression might better explain this relationship. Protective factors included the percentage of male population, perhaps due to reporting bias.
For Chikungunya, risk factors include percentage afro-Colombian population in a neighborhood (IRR = 6.1 in model 3, 10.5 in model 4, and 4.4 in model 5), average rain (IRR = 1.1-1.2 in all models), and percentage of male population (IRR = 803.6 – 2368.6). Percentage African and male changed with the random intercepts model suggesting that a geographically weighted regression would be a better choice to model this relationship. Protective factors included the 1 month lagged temperature anomaly.

For Zika, risk factors include the lagged temperature anomalies (IRR = 15.2 in the global model and decreasing to 2.3 and 2.7 in the longitudinal and random intercepts model respectively, perhaps related to timing of disease importation or alternatively that the time of disease importation was a function of temperature), percentage afro-Colombian population in a neighborhood (IRR = 2.5 in the global model, 9.04 in the longitudinal model, and 2.4 in the random intercepts model). Protective factors included average rain (0.8 - 0.9 in all models) and percentage of male population (0.7 in the global model and 0.0 in the longitudinal and random intercepts model).

Factors expected to be important risk factors that failed to show significant risk difference were: service index anomalies which were expected to be related to water storage and use with sewage as a proxy and income with energy as a proxy, social strata as a proxy of wealth and health care access, distance to canal as it was mentioned often in neighborhood interviews as the largest mosquito breeding site, and population density as a proxy for disease transmission risk. These factors may be important as interaction effects but were not explored further here.

The rain and temperature effects seen in the chikungunya and Zika models may be an artifact of the outbreak timing or the outbreak may have coincided with the rain and temperature changes, or a combination of both. Since chikungunya was introduced and quickly transmitted starting in January 2015 of an el Niño year, temperatures steadily increased and rain decreased as the outbreak progressed. Similarly, Zika was imported in January 2015 and only 4 months of data are available here.
Therefore, these coefficients cannot be interpreted without considering the temporality and limited seasonal data available for chikungunya and Zika.

Table 3: Series of models for each disease outcome. 1. global stepwise AIC Poisson regression with robust standard errors (SWAIC); 2. stepwise Poisson model (exiting p value = 0.1 and entering value =0.05), robust standard errors (stepwise); 3. global negative binomial regression with robust standard errors (nb global); 4. longitudinal general estimating equation with a negative binomial distribution and a log link and second order autoregressive correlation matrix and robust standard errors (xt nb); And 5. longitudinal negative binomial regression with random intercepts by neighborhoods (xt nb random).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain lag 1</td>
<td>1.0**(4.1)</td>
<td>1.0***(-4.0)</td>
<td>1.0(-2.1)</td>
<td>1.0(-0.7)</td>
<td>1.0**(2.15)</td>
</tr>
<tr>
<td>Lag 1 z Temp</td>
<td>1.06**(2.78)</td>
<td>1.06*(2.79)</td>
<td>1.08***(-3.99)</td>
<td>0.91***(-6.10)</td>
<td>0.96**(3.13)</td>
</tr>
<tr>
<td>Afro-Colombian</td>
<td>31.6***(-15.2)</td>
<td>17.1***(-19.5)</td>
<td>11.7***(-13.1)</td>
<td>18.4***(-6.5)</td>
<td>1.5(-1.4)</td>
</tr>
<tr>
<td>Rain</td>
<td>1.1***(-6.2)</td>
<td>1.1***(-6.4)</td>
<td>1.1***(-6.4)</td>
<td>1.1***(-11.4)</td>
<td>1.0***(-7.9)</td>
</tr>
<tr>
<td>Anomaly services index</td>
<td>1.0***(-4.6)</td>
<td>1.0(0.4)</td>
<td>1.0(-0.4)</td>
<td>1.0(0.8)</td>
<td></td>
</tr>
<tr>
<td>Strata</td>
<td>1.0(-0.3)</td>
<td>1.1**(3.1)</td>
<td>1.0(0.5)</td>
<td>1.0(-0.2)</td>
<td></td>
</tr>
<tr>
<td>Home %</td>
<td>0.0***(-7.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male %</td>
<td>0.0***(-11.8)</td>
<td>0.0***(-7.5)</td>
<td>0.0***(-7.2)</td>
<td>0.0***(-5.8)</td>
<td>0.0***(-15.7)</td>
</tr>
<tr>
<td>Single %</td>
<td>0.4(-0.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability %</td>
<td>0.1(-2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy %</td>
<td>1.5(0.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anomaly educ. index</td>
<td>1.0***(-5.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface area</td>
<td>1.0***(-5.8)</td>
<td>1.0***(-12.5)</td>
<td>1.0(2.1)</td>
<td>1.0***(-4.5)</td>
<td>1.0***(-3.5)</td>
</tr>
<tr>
<td>Population density</td>
<td>1.0(1.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to canal (m)</td>
<td>1.0***(-3.0)</td>
<td>1.0***(-3.1)</td>
<td>1.0(-1.7)</td>
<td>1.0(-0.1)</td>
<td>1.0(-1.7)</td>
</tr>
<tr>
<td>Population</td>
<td>1.0***(-4.5)</td>
<td>1.0***(-14.1)</td>
<td>1.0***(-6.1)</td>
<td>1.0(1.9)</td>
<td>1.0(-0.7)</td>
</tr>
</tbody>
</table>

N= 6441, Exponentiated coefficients interpreted as incidence risk ratios (IRR); z statistics in parentheses

* p < 0.05, ** p < 0.01, *** p < 0.001
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain lag 1</td>
<td>1.0''(-3.0)</td>
<td>1.0''(-3.1)</td>
<td>1.0(0.8)</td>
<td>1.0(2.1)</td>
<td>1.0(-2.0)</td>
</tr>
<tr>
<td>Lag 1 z Temp</td>
<td>0.8''''(-12.2)</td>
<td>0.8''''(-12.4)</td>
<td>0.8''''(-10.1)</td>
<td>0.8''''(-9.7)</td>
<td>0.8''''(-9.0)</td>
</tr>
<tr>
<td>Afro</td>
<td>12.5''''(7.3)</td>
<td>9.1''''(8.3)</td>
<td>6.1''''(5.9)</td>
<td>10.5''''(5.6)</td>
<td>4.4''''(3.8)</td>
</tr>
<tr>
<td>Rain</td>
<td>1.2''''(7.2)</td>
<td>1.2''''(7.4)</td>
<td>1.2''''(8.5)</td>
<td>1.2''''(12.7)</td>
<td>1.1''''(7.2)</td>
</tr>
<tr>
<td>Anomaly services index</td>
<td>1.0(-0.6)</td>
<td>1.0''''(10.3)</td>
<td>1.0''''(3.6)</td>
<td>1.0''''(3.5)</td>
<td>1.0''''(2.7)</td>
</tr>
<tr>
<td>Strata</td>
<td>1.1(1.9)</td>
<td>1.1(1.9)</td>
<td>1.1''''(3.2)</td>
<td>1.0(0.8)</td>
<td>1.0(0.2)</td>
</tr>
<tr>
<td>Home %</td>
<td>0.04(-1.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male %</td>
<td>67.5''''(7.9)</td>
<td>78.5''''(9.3)</td>
<td>803.6''''(12.0)</td>
<td>1953.1''''(11.9)</td>
<td>2368.6''''(15.2)</td>
</tr>
<tr>
<td>Single %</td>
<td>0.4(-0.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability %</td>
<td>0.0(-1.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy %</td>
<td>1.6(0.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anomaly educ. index</td>
<td>1.0(2.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface area</td>
<td>1.0(-1.5)</td>
<td>1.0''''(8.8)</td>
<td>1.0''''(-3.7)</td>
<td>1.0''''(4.9)</td>
<td>1.0''''(3.7)</td>
</tr>
<tr>
<td>Population density</td>
<td>1.0(0.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to canal (m)</td>
<td>1.0(-0.3)</td>
<td>1.0(-0.3)</td>
<td>1.0(0.4)</td>
<td>1.0(0.6)</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>1.0(-1.6)</td>
<td>1.0(0.7)</td>
<td>1.0(-1.3)</td>
<td>1.0(-1.3)</td>
<td></td>
</tr>
</tbody>
</table>

N= 6441, Exponentiated coefficients interpreted as incidence risk ratios (IRR); z statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001
<table>
<thead>
<tr>
<th>Zika</th>
<th>1. SWAIC</th>
<th>2. Stepwise</th>
<th>3. Nb global</th>
<th>4. Xt nb</th>
<th>5. Xt nb random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain lag 1</td>
<td>0.9***(-3.4)</td>
<td>0.9***(-3.3)</td>
<td>1.0 (1.8)</td>
<td>1.0 (-1.5)</td>
<td>1.0 (1.0)</td>
</tr>
<tr>
<td>Lag 1 z Temp</td>
<td>2.7*** (39.2)</td>
<td>2.7*** (39.4)</td>
<td>15.8*** (31.9)</td>
<td>2.3*** (48.4)</td>
<td>2.7*** (30.8)</td>
</tr>
<tr>
<td>Afro</td>
<td>21.1*** (5.0)</td>
<td>6.7*** (5.0)</td>
<td>2.51 (2.3)</td>
<td>9.04*** (5.3)</td>
<td>2.4 (2.4)</td>
</tr>
<tr>
<td>Rain</td>
<td>0.9***(-3.6)</td>
<td>0.9***(-3.6)</td>
<td>0.8***(-9.1)</td>
<td>0.9***(-7.7)</td>
<td>0.9***(-4.6)</td>
</tr>
<tr>
<td>Anomaly services index</td>
<td>1.0*(-2.4)</td>
<td>1.0 (-0.4)</td>
<td>1.0 (0.9)</td>
<td>1.0 (0.5)</td>
<td></td>
</tr>
<tr>
<td>Strata</td>
<td>1.0 (0.2)</td>
<td>1.1 (1.7)</td>
<td>1.1 (1.3)</td>
<td>1.0 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Home %</td>
<td>0.0*(-2.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male %</td>
<td>0.0***(-3.3)</td>
<td>0.0 (-1.8)</td>
<td>0.7 (-0.2)</td>
<td>0.0*(-2.5)</td>
<td>0.0***(-2.7)</td>
</tr>
<tr>
<td>Single %</td>
<td>0.9 (-0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability %</td>
<td>0.0 (-2.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy %</td>
<td>5.9 (0.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anomaly educ. index</td>
<td>1.0*(-2.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface area</td>
<td>1.0 (0.6)</td>
<td>1.0*** (4.4)</td>
<td>1.0*(-2.1)</td>
<td>1.0*** (3.5)</td>
<td>1.0*(-2.0)</td>
</tr>
<tr>
<td>Population density</td>
<td>1.0 (0.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to canal (m)</td>
<td>1.0 (0.0)</td>
<td>1.0 (-0.9)</td>
<td>1.0 (0.5)</td>
<td>1.0 (0.5)</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>1.0*(-2.0)</td>
<td>1.0*** (4.7)</td>
<td>1.0*(-2.3)</td>
<td>1.0 (0.6)</td>
<td>1.0 (0.3)</td>
</tr>
</tbody>
</table>

N= 6441, Exponentiated coefficients interpreted as incidence risk ratios (IRR); z statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

The longitudinal models (Table 3 models 4-5) estimated and interpreted here do not account for geographical variation in risk factors, especially those related to vector ecology, human activities, and urban infrastructure (unplanned urbanization, drainage systems, sanitation services, altitude, ground slope, land use, etc. (Figure 8)).

Rain was on average heavier in the east with a range of 10.2 – 1.7 mm (Figure 8 B). Distinct pockets contained higher standardized mean male proportions, especially those neighborhoods containing military bases (Figure 8 C). The distance to the canal ranged from 0 to 1,344 meters and was geographically clustered (Figure 8 E). The city services index anomalies were higher in Agua Blanca, Terron Colorado, and Ciudad Floralia (Figure 8 F). Agua Blanca contained the most people but Ciudad Floralia the highest density (Figure 8 G). African-Colombians ranged from 3.4-70.7% of the neighborhood population with a higher density in the west, especially Agua Blanca (Figure 8 D). The social strata ranged from 1-6 with higher social strata clustered in the eastern and southern portions of the city, and lower social strata in the west and northwest (Figure 8 H).
Geographically weighted Regression

The geographically weighted regressions were estimated for each disease using a Poisson distribution. There was no evidence of spatial autocorrelation in the residuals as tested by Moran’s I for spatial autocorrelation (Dengue: Z-score = 1.6; p-value = 0.1; Chikungunya: Z-score = 0.78, p-value = 0.4; Zika Z-score = 0.9, p-value = 0.3). Later, the Poisson distribution was compared to a quasi-poison and negative binomial distribution and find no difference in estimates but inflated standard errors so the simpler Poisson models are presented here in Figure 9 through Figure 11.

Each map represents one coefficient in the geographically weighted Poisson model. The darker red colors represent a greater risk factors by neighborhood and the darker blue colors represent a greater protective factor by neighborhood. The standard error of each estimate is represented by contoured lines around the neighborhood with thicker lines representing greater standard error.
Chikungunya

The GWR Poisson model predicted 55.4% of variation in chikungunya in Cali with a range of 41-67% (Figure 9 I) compared to 44.0% $R^2$ in the global model. Deviance/DOF decreased by 12.9 ($p <0.0001$) (Table 4). The model predicted a range of 0.2 to 47 cases per neighborhood with higher predictions clustered where expected in Agua Blanca and Ciudad Floralia. The residuals ranged from -23 to 35.

Table 4: AIC comparing global and GWR Poisson models Chikungunya

<table>
<thead>
<tr>
<th>Source</th>
<th>Deviance</th>
<th>DOF</th>
<th>Deviance/DOF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global model</td>
<td>1371.896</td>
<td>328.000</td>
<td>4.183</td>
</tr>
<tr>
<td>GWR model</td>
<td>1093.526</td>
<td>306.463</td>
<td>3.568</td>
</tr>
<tr>
<td>Difference</td>
<td>278.370</td>
<td>21.537</td>
<td>12.925</td>
</tr>
</tbody>
</table>

Percentage male population IRR ranged from protective in the central region at IRR = 0.6 with an estimated 40% fewer cases for each percentage point increase in male population to a great risk factor in the south east with estimated IRR of 238. Notably, a military barracks and hospital are located in this region and heavily male populated. Percentage Afro-Colombian ranged from IRR = 0.5 – 2.7, acting as a protective factor in Agua Blanca and a risk factor in the northwestern regions of the city (Figure 9 G).

Social strata seemed to act as a low-grade risk factor (IRR = 1-1.4) with clustering in the gallery city center. This area draws farmers from rural areas who are selling their goods, day-workers, and local shoppers who come for lower prices and sell in smaller local stores. This could also be related to the prison located in this region although this hypothesis cannot be tested here as incarcerated individuals were excluded from the study. This area is also strata 3 compared to its neighbors which are strata 1-2. The high income eastern mountainous regions are also pockets of wealth compared to lower income neighbors, who may be more likely to report disease (Figure 9 H).

Lagged temperature anomalies seemed to increase the risk in Agua Blanca by 213-fold; rain was protective (0.2-0.5), especially in Agua Blanca and lagged rain was a risk factor (1.2-3.5), especially in the east. Again, this anomaly could be an artifact of timing, when the outbreak hit this region, and not the seasonality of the disease itself (Figure 9 C, J, K).
Dengue

The GWR Poisson model predicted 73.1% of variation in dengue in Cali with a range of 60-78% (Figure 10 A) compared to 61.0% R2 in the global model. Deviance/DOF decreased by a 109.5 (p <0.0001) (Table 5).

Chikungunya

The GWR Poisson model predicted 55.4% of variation in chikungunya in Cali with a range of 41-67% (Figure 9 I) compared to 44.0% R2 in the global model. Deviance/DOF decreased by 12.9 (p <0.0001) (Table 4). The model predicted a range of 0.2 to 47 cases per neighborhood with higher predictions clustered where expected in Agua Blanca and Ciudad Floralía. The residuals ranged from -23 to 35 (Table 4). The model predicted a range of 1 to 550 cases per neighborhood with higher predictions clustered where expected in Agua Blanca and Ciudad Floralía. The residuals ranged from -353 to 282. Notably, Lili was underestimated by 189, El Morichal 351, Mojica by 251, and Sector Laguna del Pondaje by 235. Ciudad Cordoba was overestimated by 101 and Alfonso Bonilla Aragón by 116.

Table 5: AIC comparing global and GWR Poisson models Dengue

<table>
<thead>
<tr>
<th>Source</th>
<th>Deviance</th>
<th>DOF</th>
<th>Deviance/DOF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global model</td>
<td>9235.134</td>
<td>328.000</td>
<td>28.156</td>
</tr>
<tr>
<td>GWR model</td>
<td>6343.461</td>
<td>301.600</td>
<td>21.033</td>
</tr>
<tr>
<td>Difference</td>
<td>2891.673</td>
<td>26.400</td>
<td>109.534</td>
</tr>
</tbody>
</table>

Percentage male population IRR ranged from a great risk factor in the south-east with estimated IRR of 7.5 to protective with IRR = 0.0 (Figure 10H). Notably, a military barracks and hospital are located in the at-risk region, and heavily male populated. Percentage Afro-Colombian ranged from IRR = 0.6 – 8.0, acting as a protective factor in Agua Blanca and a risk factor in the northwestern regions of the city (Figure 10 E). Social strata appeared to act as a low-grade risk factor (IRR = 1.3) in the gallery city center and the eastern mountainous regions, and as a protective factor in the south-east near the barracks (Figure 10 K).
Lagged temperature anomalies seemed to increase the risk in Agua Blanca by 390-fold; rain was protective (0.2-0.5), especially in Agua Blanca, and lagged rain was a risk factor (1.2-4.2), especially in the east. Again, this could be an artifact of timing, when the outbreak hit this region, and not the seasonality of the disease itself (Figure 10 C-D, L).

Zika

The GWR Poisson model predicted 61.5% of variation in Zika in Cali with a range of 51-69% (Figure 11 K) compared to 47.0% R2 in the global model. Deviance/DOF decreased by 26.6 (Table 6). The model predicted a range of 0.3-70.7 cases per neighborhood with higher predictions clustered where expected in Agua Blanca and Ciudad Floralia (Figure 11 J). The residuals ranged from -48 - 58. Notably, El Morichal was underestimated by 58, Lili by 34, and Mojica by 25. Ciudad Cordoba was overestimated by 21 and Manuela Beltrán by 48.

<table>
<thead>
<tr>
<th>Source</th>
<th>Deviance</th>
<th>DOF</th>
<th>Deviance/DOF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global model</td>
<td>1725.820</td>
<td>328.000</td>
<td>5.262</td>
</tr>
<tr>
<td>GWR model</td>
<td>1251.634</td>
<td>310.187</td>
<td>4.035</td>
</tr>
<tr>
<td>Difference</td>
<td>474.186</td>
<td>17.813</td>
<td>26.621</td>
</tr>
</tbody>
</table>

Table 6: AIC comparing global and GWR Poisson models Zika

Percentage male population IRR ranged from protective in the central and northeastern region at IRR = 0.1 with an estimated 90% fewer cases for each percentage point increase in male population to a great risk factor in the south-east with estimated IRR of 76.6. Percentage Afro-Colombian ranged from IRR = 0.6 – 13.7, acting as a protective factor for most of the city and a risk factor in El Refugio and Camino Real- Joaquin Borrero Sinisterra (Figure 11 E). Social strata seemed to act as a low-grade risk factor (IRR = 1-1.5) in the same regions as Dengue and chikungunya (Figure 11 D).

Lagged temperature anomalies seemed to increase the risk in Agua Blanca by 92.9-fold; rain was protective (0.2-0.5), especially in Agua Blanca, and lagged rain was a risk factor (1.2-2.8), especially in the east and south-east (Figure 11 A, H, L).
Factors that were not significant in any of the models included surface area, distance to canal, population, and service anomalies.

Figure 9: Chikungunya coefficients and standard error, January 2015 - April, 2016.
Figure 10: Dengue coefficients and standard error, October 2014 - April, 2016
We find that percentage population Afro-Colombian and male and social strata are significant factors in predicting disease outcome.
El Morichal, Lili, Mojica, Ciudad Córdoba, Manuela Beltrán, Sector Laguna del Pondaje, and Alfonso Bonilla Aragón were outliers. This may be either related to reporting bias, or may point to some interesting differences by region which should be further explored.

The percent of deviance explained ranged from 3.3 to 73.2% with the GWR for dengue performing the best (73.2%) followed by the GWR for Zika (61.6%) and finally the GWR for Chikungunya (55.3%). The AICs for chikungunya ranged from 1,151 to 9,243 with the GWR performing the best. The AICs for dengue ranged from 6,401 to 26,328 with the GWR performing the best. The AICs for Zika ranged from 1,309.3 to 6,286.2 with the GWR performing the best. There was no evidence of spatial autocorrelation in the residuals of the GWR (Moran’s I p-value = 0.4, 0.1, and 0.3 for Chikungunya, Dengue, and Zika respectively). All GWR’s were tested at 62 meters. The Wald chi square test was significant at p<0.0001 in all model’s tested (Table 7).

Table 7: comparison of models

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Chikungunya</th>
<th>Dengue</th>
<th>Zika</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Global XT</td>
<td>GWR</td>
<td>Random</td>
</tr>
<tr>
<td>DoF</td>
<td>12</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>AIC</td>
<td>9243.5</td>
<td>6401.2</td>
<td>9133.2</td>
</tr>
<tr>
<td>R²</td>
<td>11.0</td>
<td>26.4</td>
<td>55.3</td>
</tr>
<tr>
<td>Moran’s I</td>
<td>0.01, Z = 0.8, p = 0.4</td>
<td>0.02 (0.4)</td>
<td>0.0</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>62.0</td>
<td>62.0</td>
<td>62.0</td>
</tr>
<tr>
<td>Wald chi²</td>
<td>533.5*</td>
<td>1306.7*</td>
<td>1124.7*</td>
</tr>
</tbody>
</table>

* pseudo r²; †significant at less than 0.0001; Observations/Groups= 19/339

The best model was the GWR which allowed the risk factors to vary spatially. Longitudinal model might perform better during non-outbreak periods. However, all the models produce the same conclusion thus these estimates are robust to model selection and parameter choices.

Comparison to others

Another group reported that dengue from the 2011 outbreak was associated at the neighborhood level associated with tires shops, nurseries, sewage system, and SES. We find that SES is
important but do not find a significant relationship between sewage system and disease. We have added to this model by including monthly precipitation and temperature data, and longitudinal Poisson and negative binomial model, which shows no evidence of spatial autocorrelation, and 2014-2016 data of chikungunya, dengue, and Zika. The same group compared statistical and expert based models of risk of dengue in Cali (2011 data) and found that risk is clustered spatially in agreement with our results and that risk factors include young age structures, low education, and employment or housework. A model of all Colombian dengue cases reported from 2007-2010 found a nonlinear effect of average monthly precipitation and an outbreak in 2010. Another Colombian model found that outbreaks occur during hot (18°C and 32°C) and dry (perhaps related to water storage) periods and can predict dengue outbreaks 2 weeks to 6 months in advance.

Phenomenological Models of Zika in Antioquia, Colombia during January-April 2016 found that the reproduction number changed as the outbreak progressed (from 10.3 in the first generation to 2.2 in the second disease generation, and warns that these models are highly sensitive to model assumptions and parameters. Majumder et al collected Zika data from health map and compared it to data from weekly Instituto Nacional de Salud (INS) epidemiological bulletin publications and used The Incidence Decay and Exponential Adjustment (IDEA) model to model R0 and Robs for both and found HealthMap R0 range 1.42-3.83 Robs range 1.42-2.30 INS R0 range 2.34-8.32 Robs range 1.60-3.31 and concluded that in the lack of traditional surveillance data, non-traditional digital data can provide suitable estimates with larger ranges. Another non-traditional form of data came from big hourly cellphone data in china to map intra-urban transmission risk of dengue fever and quantified uncertainty of importation risk and used the Spatio-temporally resolved programs for city-wide dengue control.

We agree with Kienberger et al in that data should be made available to share among collaborators, policy makers, and community health workers in online fora. An open source platform, DengueME, is supporting the generation of transmission models and “supports compartmental and
individual-based models, implemented over a GIS database, that represent *Aedes aegypti* population dynamics, human demography, human mobility, urban landscape and dengue transmission mediated by human and mosquito encounters". However, the author looked for both models and could not find the open-source models/software/data.

The strengths of our models are the robust estimates across model types, the inclusion of spatial and temporal factors and climate data. We were also able to validate our models across various disease outcomes. Limitation of this study include that the census data used here has not been updated since 2005. Also, with so few months of Zika and chikungunya data, the seasonality cannot be interpreted as was done with dengue. Despite multiple iterations and expert insight, some cases were not able to be georeferenced and were excluded (1,688/26,985 = 6.3%). Finally, there may be variability within the neighborhood level that was not captured with these neighborhood level models.

**Conclusions**

We predicted up to 78% of variability in disease outcome by neighborhood risk including the district of Agua Blanca. Some areas expected to be at high risk (e.g. Terron Colorado) but were not. Others were at higher risk than expected (e.g. Lili), which represents a public health intervention opportunity. For example, these areas could be underreporting and should be further investigated or these could be preventing the disease better than other neighborhoods and could be areas from which public health officials and investigators can learn. From this model, the local granular level data, which might explain this variation in risk at the local level, are missing. Future studies should include seroprevalence surveys to estimate true burden of disease and continued collaborations with other mappers, MoH, and CHWs. Even the GWR model cannot account for local variability beyond the neighborhood level and the census data has not been updated since 2005. To capture finer scale variability in environmental factors related to human behavior, vector proliferation, and the interaction
between the two, further studies should be conducted at the neighborhood level focusing on areas of highest risk.
Chapter 4:

Aim 3

Synopsis

The risk factors for vector-borne diseases (poverty, trash, water accumulation, limited access to health care, vector abundance) are known. Locals who know Cali and its neighborhoods know that these risks exist in varying degrees across the city. It can be argued that if these factors are not measured and reported, the concomitant risk will not change. Here, these risks are measured and factors are associated with disease presence at the neighborhood level. These data provide the bases to create community education and intervention plans, and provide evidence to ask for additional support. This brings public health, environmental health, vector management, social justice, and policy together.134,205

Manuscript III:

Assessing Risks from Dengue, Chikungunya and Zika at a sub-neighborhood scale of Cali, Colombia using spatial video, spatial video geonarratives, and the community perspective

Abstract (350)

Background: This is a study to describe the geographic relationship between the dengue incidence with the outbreak of chikungunya and Zika in Cali, Colombia during 2014-2016. Cali exhibits Hypoendemic dengue transmission year-round and experienced outbreaks of chikungunya and Zika in 2015 and 2016 respectively. We include here local expertise from community members, ministry of health, and formal and informal community health workers in the form of geospatial narratives.
Methods: Using secondary case data from SIVIGILA, the cases were georeferenced. We performed kernel density and hotspot analysis on the data to identify areas of interest for mapping by geospatial narratives. Those geospatial narratives were translated, transcribed and mapped over hot spot areas to give local context to the hotspots. Water related risks were mapped along with disease burden.

Results: We find that dengue and Zika exhibit similar geospatial patterns while chikungunya was more widespread. Local sentiment is that all three diseases are one risk, mosquitoes are a nuisance related to the canal, poverty, and invasion, and the MoH should do more fumigation but with less toxic methods. The Ministry of health participated in interviews and was seen cleaning storm drains and fumigating during data collection periods. Future work will include continued collaboration with local collaborators to socialize and validate results. These local disease risks should also be the focus of local intervention. Water risk overlapped local disease risk.

Conclusions: Here, the first field validation of geospatial narratives for vector-borne diseases in South America are presented. These findings are relevant to local epidemiology as these allow the local resources to be optimized for vector control. These methods can also be used in other areas to guide MOH in resource management. The risks identified contribute to local quality of life and should be the focus of public health and community intervention.

Introduction

Zika in the Americas during 2016 has received considerable coverage in popular media outlets. Yet many of the locations vulnerable to Zika are also endemic with two other diseases carried by the same mosquito and which are, arguably, more of a public health concern; Dengue and Chikungunya. Of these, Dengue has generated a large spatial research body, mostly focused on various components of the spatial and spatial temporal patterns of either mosquito intensity or human case data. Here, by focusing on the lesser covered topics of sub-neighborhood granularity and mixed-method contextual
insight, this literature has been expanded. The need for a more granular understanding at a sub
neighborhood scale that incorporates specific features such as houses, streets, standing water and other
disease promoting situations has been identified and was previously acknowledged as an important next
direction \(^{64,108}\). The addition of human context or insight is understood to be important as it has been
identified in other spatial research areas \(^{81,109,110,111}\). The combination of these can help explain why
spatial variation occurs between neighborhoods of an endemic city. For example, how does local human
behavior cause / mitigate disease risk \(^{111,118}\), and what is the on-the-ground perception of disease risk
from within surveillance “hot spots”? For this study, two new epidemiological techniques are further
developed, which can be used to address these topics, and in so doing, enrich an already vibrant
research agenda focused on dengue in Cali, Colombia.

A core component of most spatial research of mosquito vectored disease involves the
identification and explanation of geographic patterns of mosquitoes, and especially disease carrying
mosquitoes, or human surveillance data \(^{62,111,112,113,114}\). Causation and correlation with these patterns
involves the incorporation of various spatial layers of which environmental (including moisture and
vegetation), built environment (infrastructure), climatic (micro and macro), entomological, and human
(density, social, behavioral, political and disease surveillance) are the most common \(^{109,118}\). Adding
complexity to these studies is an issue of geographic scale and the realization that the pattern-
cause/correlation nexus will vary geographically; between countries, regions of countries, cities, and
even within cities \(^{114}\). Though there are still accepted universal risk factors for dengue, such as human
density, various measures of economic hardship \(^{62,81}\), and proximity to urban environmental “risk”
features such as canals and ditches, even these will vary in importance across the same urban space due
to variations in the “attributes” of each, such as how the micro environments of the same canal will
change risks in proximity. Adding spatial and temporal dynamism to these patterns are intervention
methods such as mosquito control and disease education strategies \(^{115,116,117}\). Spatial research at a more
granular *scale of intervention*, especially in combination with on-the-ground perceptions and insights can help reveal these types of nuances that will lead to the more general map of risk. In this paper, this granular scale is the focus by showing how street level environmental data extracted from spatial video (SV) and Google Street View (GSV), and mapped expert insight captured by spatial video geonarratives (SVG) can be used at any location to better understand disease risk.

*The Dengue situation in Cali, Colombia*

Dengue is a viral disease spread by mosquitoes of the aedine family, especially *Aedes aegypti* in the Americas. Control efforts had been relatively successful during the middle of the last century but a variety of challenges to and delicacies in control beginning in the 1970s (particularly those experiencing considerable conflict and violence) allowed the reemergence of the disease.

The metropolitan City of Santiago de Cali, Colombia (hereafter referred to as Cali) covers 560.3 Km², is located at 3°27'26"N and 76°31'42"W, is 1,070 meters above sea level, 3 hours from the pacific coast of Central Colombia in the department of Valle de Cauca with a population of 2,181,317 in 2010. The major ethnic groups are Afro-Colombian (26%), Caucasian and mixed (73.3%), and indigenous groups (0.5%). The national median GNI per capita was US$7,560 with 26.1% of the population in Cali living in poverty. The primary cause of death is homicide (including traffic accidents) followed by cerebrovascular events. The median life expectancy in 2010 was 74. As previously described, between September 2014 and September 2015, there was an outbreak of Chikungunya in Cali. The outbreak has ended with continued endemic transmission. Chikungunya transmission occurs simultaneously with dengue by the same vector in the same target populations. Both are being controlled as a stable endemic disease.

---

We will use examples taken from these expert narratives throughout the paper to illustrate the various points being raised.
Cali, Colombia provides an excellent study location to show the importance of a granular perspective on Dengue. It has long been recognized as having a mosquito vectored disease problem which has occurred not only because of similar regional trends across the Americas, but also because of the violence experienced in Colombia which disrupted health protection. The city has considerable variation in neighborhood type, and a rich history of (dengue related) research which has considered the various aspects of disease presence, and the associated risk factors in an attempt to understand disease patterns, and therefore potential interventions, in such a complex urban environment 62, 63, 64, 65, 66.

Spatial and spatio-temporal patterns of disease, environmental and human correlates and causations, access to health care and mobility, have all be considered with regards explaining Dengue in Cali. For example, Delmelle and colleagues81 used geographically weighted regression to tease out neighborhood scale variations within the city. Their work not only included a rich array of input variables (cemeteries, 49 plant nurseries, 23 water pumping stations, 499 tire shops, 4 rivers, and green spaces) but also incorporated local expert insight in terms of what they initially should have considered. Even given their results, which validated many other similar findings with regards to the vulnerabilities posed by human density and socio economic status, they acknowledge that richer local and contextualized data are required. There are scale issues to be considered as their work was forced to aggregate to appropriate census units. Also, there are likely to be surveillance deficiencies based on some of those variables found to elsewhere to be predictors of vulnerability.

This last point is of particular concern for urban environments like Cali if dengue vulnerable neighborhoods also coincide with perceived and actual security risks 59, 119) which may limit the work of vector control in the most at risk neighborhoodsb. In our experience, many community members are

---

b S3: “The administration never takes into account the leaders. Never. Those are the ones that know where the production is of the effect that is affecting the population. They never. They prefer to spend the money on publicity and not call the leaders to look for people to go with us and show us where the hot spots are. With one hot spot that contaminates one quarter of the city of Cali. They never do that. If you hear the secretary of health, I walked and went and they go to xxx, but what those that serve for? Celebrating parties to the mosquitoes. That is
aware of this bias and attempt to bridge the gap. Human case data (at least in Cali) may be skewed against economically deprived areas where clinical testing may be less frequently used due to a lack of health care access or a lack of confidence in the health care system. Additionally, from a spatial perspective, a worryingly low percentage of the disease surveillance data can be geocoded to an actual or proximate address. Even if these data are accepted as an unbiased sample (which is debatable)

ridiculous to put on TV. If you look at the seventy percent of officials [in the vector campaign] that they named, they are afraid to go the neighborhoods, the sidewalks. They're afraid.”

S3: “Principally I would offer [you] security because we can’t get a quadrant during what we can do and look at the places where the hot spot of the disease is. Where it is born. Yes, where it is born. There we have to focus on it and look at it. S2: ...my father guarantees security because people in the neighborhood respect him a lot. They do not go in between my father because they know my dad does social work in the neighborhood. Then, if you are going to go, go with him.”

S2: “For example, here, we have three nurses that are-- we do community work. Now a days, I do not practice. I dedicate my time to that, with hers. We are three and we do social hours. We recollect the drug that a lot of people, more than anything the ones from here because the ones from Calipso we have a lot of people that bring drugs to us.”

I1: “Medicine.” S2: “Yes medicine. The medicine but we do not give the medicine out to any one. We go and do brigades. We take the medicine...Medicine that we have already there. We look at the expiration date that is good. And we go do the brigades.”

d V1: “Santa Helena has a canal- a canal of rainwater. What happens is that as Santa Helena is so commercial, if there are cases there, everybody goes to their house and in their house, they look for remedy and they don’t report it maybe, no?”

CHW1: “Look, we had on one occasion, we were, how many groups did we have in that time, we had about 7 groups of community mothers, and we counted almost 11 cases of chikungunya. Yes? In any committee that they did, ’ah no, in that group there are this many incapacitated for chikungunya.’ yes? but I tell you because it was the mothers who said ‘I have chikungunya.’ Because the medical diagnosis of chikungunya didn’t reach us.”

S1: “And well, that is what we need, thank God, if you would see how [the hospital] fills up. And it maintains that way... They [the health center] took someone and they did not want to serve them... Look, I haven’t been there a lot, but people go, I did not go—It has been a while since I went ...You know, I went there once and I remember I had dengue. I was burning with fever and with a bad headache with shivers but I did not have any outbreaks or anything. My husband lost his job and did not have EPS. Then I went there and they told me I needed an identification and I told them my husband’s situation—I had [insurance], but since my husband lost his job I don’t. Sign up first they said, and I told them that sure, my sickness will wait while my husband comes and gets settled, or while I do the SISBEN...And then how are we here. That’s not how it is. That’s why it is called health center because you have to—Hi, Lily, come and tell us the experience you had with the health center when— I told them you had Chikungunya fifteen days ago and you went to the health center and they said... “S2: They told me they could not tend to me because I had to go to the church there. I told them they always tend to me there. They told me that—I have the SISBEN there and I had to go.”
given the problems identified with poverty), both problems of context and interpretation remain, and for granular data to be used in any explanatory analyses. For example, a neighborhood may present as a hotspot – but why? More specifically, why does the east side of the canal appear to be more disease prone than the west? This hints at causations may be more complex than a simple proximity to a canal but may involve directionality. The same variation occurs across the course of the canal as the flow of water may be linked to rain runoff, but micro variations will occur that result in water stagnation such as engineering work, increased trash (especially around informal housing) or edge vegetation all of which can increase risk. While techniques such as Geographically Weighted Regression (GWR) have been used to gain insight into such neighborhood heterogeneity, these are still limited by the coarse nature of data input which have to be standardized, most frequently to a census enumeration unit. Yet all canals, all parks, and neighborhoods of poverty will vary geographically in terms of localized risk. It could easily be that a single section of canal (for the reason identified above), one park (or park corner where dumping occurs), and variations in the cultural and political nature of trash accumulation generate a much broader “hotspot”. As mentioned, data ecology problems exist with census aggregations, both in terms of surveillance data having to be aggregated to these units, or from the general lack of

---

6 V1: “Look how the canal is. It is completely stagnant. You know why? Because of the construction more ahead. And look at the larvae how you can see them there.” I1: “All that movement is larva?” V1: “Yes. Yes, look at how they are moving in the water. Is because they are fixing something more ahead and the machine is inside in the middle.” I1: “But for a lot of months?” V1: “Yes.” I1: “Is it like that for a lot of time?” V1: “Yes because the other time we came those borders in the wall weren’t there. They were making them, of the canal.”

S1: “Yes, we here in Yira Castro, by the Troncal, we have a canal that we have suffered from because in the canal you get close at four or five of the afternoon and see thousands and thousands of mosquitoes.” I1: “So you told us that the mosquito appear during five thirty in the afternoon in those areas where they are constructing, fixing the sewers because residual waters gets accumulated and towards the side where the canal where they are fixing it. What solutions do you think they are for the canal and for when they are constructing or fixing the way... S1: Well I as committee-- I am apart of the committee for plan making here in Comuna Trece. We came to agreement so they could cover the canal.”

I1: “The large parts of the cases are in these sectors.” S1: “Well, I blame the canals a lot. Because that canal, I would like for us to pass in the truck because it is full of dirty water. Go look at it and take a picture and you will see the amount of mosquitoes there. They are thousands. They are small and they are big. And you see some mosquitoes and their legs are white. Those white legs.”

64
appropriateness as the research moves further away from the census collection date, or having to be forced into using environmental layers that exist at those scales. GWR is certainly a useful tool to provide a first analysis, but it should be considered as the first step, identifying sub neighborhood spaces for more detailed investigation.

Potentially as important as having more granular data is the addition of local insight to contextualize risk. For example, in their study on mosquito control in a medium sized US city, Curtis et al used SVG to capture institutional knowledge from local vector control experts; spatially specific insights from those who had worked for a number of years in the area and who had amassed valuable information on topics such as which locations cause mosquito outbreaks, which species were involved, what control measures had been (successfully) used and what were the proximate human diseases. SVG were used to map locations such as basements of abandoned buildings, and even areas with control challenges such as a high number of potted plants. In their pilot study a SVG baseline was established for potential future reference. The resulting paper also raised the possibility of a similar approach being applied to a more severe disease environment, one endemic with disease and with fluctuations in intensity, such as Cali, Colombia.

**Spatial Video and Spatial Video Geonarratives**

Spatial video (SV) has been used to map various types of challenging and data poor environments. In Haiti, environmental layers (standing water, trash and animals) were mapped at a granular scale to help explain urban cholera patterns and then support epidemiological field testing. The basic data layer creation involves watching spatially encoded video and then digitizing risk factors into either Google Earth or a geographic information system (GIS). While this does not replace

---

\(^{1}\) In Demelle and colleagues study tire stores were included as a measure of risk, but as the authors note, no individual tire store assessment is included and this will obviously lead to an inappropriate categorization of risk as *not all tires stores are the same*, either in terms of volume or cleanliness.
more traditional epidemiological approaches, it can complement them with local spatial detail, especially in spatial-data-poor environments. The SVG method is SV with an audio recorder added to capture “expert” insight during data collection. Software written at Kent State is then used to produce mapped layers from these transcribed commentaries. If there are enough SVGs (the depth scenario), key words can be queried out as a point layer and used as input for more traditional spatial analysis \(^8\).

Alternatively, a single expert is used to cover multiple spaces (the breadth scenario) and to provide institutional insight where none exists \(^1\). For an investigation into dengue, these narratives can be used to gain further granular insight into a neighborhood or area known to be a hotspot, with the identification of specific features or even behavioral / political actions that increase vulnerability.

There are three types of insight gained from either the depth or breadth SVG: *Spatially specific*, where an exact location is identified, such as a collection of abandoned tires. These data are usually described as the vehicle passes the location, and while there is an uncertainty space with regards the geographic extent of the comment, the final point on the map is still near the actual risk. The ability to return to the video for visual validation not only helps in contextualizing the comment, but can also reduce the uncertainty of the location to an exact place. For example, Figure 12 displays a small drainage canal which is described as problematic for mosquito breeding while also being proximate to a child play space.

---

\(^8\) V1: “Look at the canal. ...” I1: “Lots of trash. Is it raining?” V1: “No, I think it’s the larva that make the water move like that. Larva, tadpoles, who knows how much, look at the other side, it looks terrible.”

V1: “Do you see why there is dengue in Calipso, look. The canal of ...look at the trash. Look, do you see. This is maintained full. Not empty.” I1: “And it is not moving at all. It is stopped.”
Spatially proximate, is a general risk space identified from a location's sights sounds and smells\(^h\) even though no precise place is identified. An example could be the description of how rainfall would often flood this part of the neighborhood which, in combination with unfinished road surfaces, would produce small pools for breeding. \textit{Spatially inspired} or \textit{fuzzy spaces}, contain information that emerges during the SVG ride about the topic under investigation without referencing spatially specific locations. For example, insights into the challenges facing vector control\(^i\), what it was like to contract dengue,

\(^h\) It should be noted that while previous research has stressed the visual side stimulating comment, actually it is all the senses working together. This is especially true with smells being an indicator of trash and stagnant organic filled water.

\(^i\) S1: “OK, look here the puddles in the road where the dengue mosquito propagates a lot.” S1: “the problem of invasions. Look here. The problems that occur in the high part accumulate in the low part of Terron Colorado and this is what more brings, more brings the mosquito. The humidity. The streets broken. There they make the puddles and begins the whole thing with the mosquito.”

\(^i\) CHW1: “this, for example. The old constructions that they don’t finish. Puddles form in the cement. Sometimes, they don’t continue.” I1: “this makes that it is more difficult to control the vector, no.” CHW1: “of course. uuhh. because they don’t climb up there to see if there are puddles of water. then there stays the vector.”
chikungunya or Zika\(^k\), or the limitations in the local disease surveillance system are all examples\(^1\).

However, even with these there can still be a geographic stimulus – for example being in an economically disadvantaged neighborhood could result in comments on vector control or disease surveillance, \textit{for that type of neighborhood}.

\textbf{Methods}

SVG were used in Cali, Colombia to capture and map sub neighborhood risk factors associated with Chikungunya, Dengue and Zika. These insights were to be gleaned from “expert” passengers who had detailed information about each neighborhood. At the same time, the SV component of each SVG would be used to digitize a granular risk map for each neighborhood. Neighborhoods would be chosen based on an understanding of the geographic variation of disease (both from existing publications and the authors own analysis of surveillance data).

Two Contour 2 cameras, designed for extreme sports usage, were mounted on the inside left and right window of a vehicle using a suction mount. The cameras were angled down to capture the conditions by the side of a road and to limit the unease of passersby who might not want to be recorded (Figure 13 a and b). Figure 13 a displays the location of the camera on the window, while Figure 13 b shows the image captured by this camera at this location. During the drive, stops would occasionally be made if the environment was not suitable for vehicles, or if added detail were needed regarding a key location (Figure 13 c). On these stops one camera was removed and hand carried to record local

\(^k\) S1: “Yes, I had it [chikungunya]. It started with itching and then an outbreak and that—I had a small fever. And as I had articulation pain because that day I started like that and the other day I went to get up to get my kid ready to school. When I went to stand up, I couldn’t because of the pain in my knee and I fell back into the bed. Then I told my son to help me stand up and he did but I walked like a robot, because I could not take care of myself. I couldn’t with the pain. I had to go around like a robot, stiff.”

\(^1\) S2: “There are a lot of people that are careless. We have the Sisben and that is excellent. But there are people that are not interested. They say ‘Oh, why am I going to go there. No, with these pains and fever to go make lines and I don’t know what else. No, I’ll stay here in my house. And so they can send me to buy Acetaminophen and I do not have anything to buy it with.’ Then that is why we go to them. Because the leaders told us, we have this problem.”

68
conditions. Figure 13 c shows the hand recording of mosquito larvae by the two individuals seen in Figure 13 b. Inside the car an audio recorder was connected (through a splitter) to the input audio jack of each camera. The expert’s commentary was primarily captured by the recorder, but his/her voice was also captured on the video. This was important for syncing the eventual transcription time stamp with those of the video and GPS.
Figure 13  a and b: Two Contour 2 cameras mounted on the inside left and right window of a vehicle- angled down to capture the conditions by the side of a road); C: hand recording of mosquito larvae by the two individuals seen in Figure B
Each ride would usually capture more than one neighborhood, though the equipment was turned off between destinations. The length of ride varied according to several factors, including size of neighborhood, accessibility within the neighborhood, and perceived safety issues. After data collection, the video was downloaded from each camera in Colombia. Upon returning to the United States the audio of each SVG was translated from Spanish and the earliest part of the transcription that could be matched to the video was used as a starting point for the word interpolation. For this to happen the same transcribed word had to be clearly heard on the video, and the GPS had to have been acquired for that video segment (sometimes GPS took longer to acquire). The software Contour Storyteller (examples are seen in Figure 12 and Figure 13) was used to view the video with an associated map, while the software was also used to extract the GPS path (as both a GPX and a CSV file). The CSV file was used to find the associated Greenwich Mean Time (GMT) as a six-digit number. The transcription (as a text file), GPS (as a CSV file), and the six digit GMT time were entered into Wordmapper, a specially written computer program created in the GIS Health and Hazards Lab at Kent State University. Output from Wordmapper includes a GIS shapefile of words, where each word is interpolated over the time range between two conversation time stamps, and a shapefile of "comments" where each conversation segment is attached to the coordinate of the beginning word. The reason for these two outputs is that these allow for a single word to be queried in the GIS (show all mentions of trash), while also spatially important conversations can be extracted around an object being discussed, which helps in terms of returning to the video for further visual investigation. A third file, a comment CSV file was also extracted for added flexibility, allowing for the potential input in other non-spatial textual analysis software, such as Nvivo.

After each neighborhood, had been mapped by spatially encoded words and geo located comments, a series of exploratory analyses were used. The word file was queried for similar risk factors (all mentions of water or water features) and these were mapped using a kernel density estimate (KDE)
surface to aid in the visual interpretation of the word points. KDE is a commonly applied technique in epidemiology, with several examples using mosquito related data, including previous research in Cali, Colombia \(^{66,81,82}\). KDE has also been previously used to map out SV and SVG \(^{83,84}\). Each KDE was classified into ten equal breaks, and these were used to determine the minimum contour value of the resulting interpolation of the KDE grid.

Each transcription was also read for general insights, and then spatially specific locations, especially those important to understanding local mosquito activity (such as a known “problem” park, or a stagnant water trench, or a discarded pile of tires). These spatial insights were identified in the GIS using the comments shapefile, and each object was buffered to 50m. Key larger features identified in the commentary, such as a canal, were digitized according to their shape and again buffered. In this way, a risk map for each neighborhood was created.

Concurrently to the SVG analysis, a few key neighborhoods known for identified during the ride as being of particular interest were digitized for visible risk factors, using the same approach previously employed in Haiti \(^{83}\). SVG was viewed on one screen while risk locations were digitized into Google Earth. Concentrations of digitized points were used to locate water risk, trash (and tire) locations, and the presence of dogs. The intensity of points capture the spatial extent of the risk, while the weight assigned approximated the depth or volume. These Google earth layers were saved as KMZs and imported into ArcGIS 10.4 where the same KDE approach as previously described was used (50m kernel, classification of ten equal breaks). As Google Street View (GSV) has also been suggested / used as a granular data source for research \(^{213,214,215}\), this was also used with the same method as described for GSV being followed. As it has been noted that one limitation of GSV is the fluctuation in time frame across the same apparent coverage \(^{216}\), all image time stamps were recorded.
Mapping disease risk

Using secondary dengue, chikungunya, and Zika case data reported between October, 2014 and April, 2016 to the national database, SIVIGiLA, cases were georeferenced using The Secretary of Health Software to identify geographical hotspots and comparative control areas, or “cold spots,” in Cali.

Maps of chikungunya, dengue, and Zika risk were constructed using ArcGIS 10.3 and its spatial analyst extension tool Getis-Ord General G statistic at 2500, 2500, and 3400 meter bandwidths respectively (chosen to optimize spatial autocorrelation analysis) and Kernel Density at 600 and 200 meter bandwidths for full extent and zoomed analyses respectively (Figure 14 & Figure 15).

In the aggregated hotspot analysis, chikungunya did not overlap Zika and dengue as expected. In the non-parametric kernel density analysis, the highest risk areas overlapped for Zika, dengue, and chikungunya.
In these hotspots, 107 geospatial interviews were conducted in 26 neighborhoods during drought (October 14 – December 1, 2015 and January 25 – February 9, 2016) and rainy season (April 9 – 16) with vector specialist, community health nurses, program leaders and local community leaders and members. Interviews from the geospatial narratives were mapped on top of hotspots to give context and key word density mapped. Interviewees sometimes gave suggestions for other community leaders to be interviewed and so the sample snow-balled. Some interviews were conducted like focus groups at home with maps and videos to be reviewed or in local community centers and some interviews were group walks to areas of interest.

Key themes included: water, violence, canals, mosquito control, and converting open canals to canals covered with parks and soccer fields.

Geonarratives

A total of 24 neighborhoods were covered using SVG across four time periods between November 2015 to April 2016 for the city of Cali, Colombia. The subject for each ride was either a mosquito control expert, a public health official or a local resident. For some rides, more than one expert was in the vehicle. Each ride began with the reading of an IRB permission sheet and gaining consent. The path of the SVG was determined by the researcher selecting the neighborhood, but the expert identifying which features were important to visit inside the area of study. In all cases the expert had a solid working knowledge of the neighborhood being studied.

\(^{m}\) It should be noted that GPS problems occurred in a few neighborhoods, while in others the amount of commentary was too short for useful risk mapping. Only the “successful” SVG are reported here.
Table 8 provides a summary of all rides, including the neighborhood, date, and the expert type being interviewed. The length of each ride is also presented in terms of the total video length, and the total of spatially coded words. The number of spatially important mentions is also tabulated; these being identified locations that potentially play some role in the mosquito situation for each neighborhood.

Finally, as an example of how the spatial word file can be investigated, a count is provided of “invasion” which is a local expression for an informal settlement. This word is chosen because these invasions were frequently mentioned as playing an important role in the dengue situation, in terms of having poor infrastructure, more trash, less mosquito control (for both political and security reasons), and insufficient access to medical facilities, and as a result questionable surveillance data\(^n \) 217. This column is

\(^n \) S3: “What I see, is that a lot of cases were presented symptomatology but no diagnosed.” S2: “But not diagnosed because people did not like it.” S3: “Simply because people would say they had articulation pain, I have fever, is dengue or they matriculated they as they said.”

CHW: “But equally from the sectors of community mothers that we had here, there were many incapacitations for chikungunya...What happened with those diagnostics- they got the incapacity and the doctor said “this is chikungunya,” but the diagnosis that they had was to discard the dengue, initially, yes, and then they were working with disability and when they came back they said it was a chikungunya for the symptoms. That is to say that they knew the symptoms of chikungunya so they told us it was a chikungunya. ‘I got a rash, my joints hurt.’ So they assumed it was a chikungunya. but a diagnosis, like a medical diagnosis, we didn't find. ...we always found to
provided as an indicator of the type of query that can be made, and any future poverty related investigation would combine all synonyms, especially as some experts preferred other terms, in some locations, to “invasion” (such as levee). The SVG rides ranged from approximately 90 minutes to 20 minutes, with the number of associated spatial words being loosely connected with trip length. One ride (Antonio Nariño) is split as the interview was broken into two segments.

Table 8: spatial video and spatial video geonarrative field collections and resulting mapped length, words, and spatial mentions.

<table>
<thead>
<tr>
<th>Neighborhood</th>
<th>Date</th>
<th>Length (mapped)</th>
<th>Total Words</th>
<th># spatial mentions</th>
<th># Invasion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfonzo Lopez 1</td>
<td>11/22/2015</td>
<td>71.40</td>
<td>2030.00</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>Antonio Nariño</td>
<td>11/26/2015</td>
<td>86.35</td>
<td>1291.00</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>Nueva Floresta 1&amp;2</td>
<td>11/24/2015</td>
<td>58.18 + 26.17</td>
<td>2332 + 393</td>
<td>29+2</td>
<td>7</td>
</tr>
<tr>
<td>Calipso</td>
<td>11/26/2015</td>
<td>32.27</td>
<td>769.00</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Lili</td>
<td>11/27/2015</td>
<td>24.30</td>
<td>962.00</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Melendez</td>
<td>11/27/2015</td>
<td>22.50</td>
<td>1436.00</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Caneo</td>
<td>12/4/2015</td>
<td>37.47</td>
<td>857.00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mariana Ramos</td>
<td>12/4/2015</td>
<td>65.10</td>
<td>1849.00</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Valle Grande</td>
<td>12/4/2015</td>
<td>63.06</td>
<td>926.00</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>Villa de Lago</td>
<td>12/4/2015</td>
<td>27.05</td>
<td>582.00</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>El Vallado</td>
<td>12/5/2015</td>
<td>36.35</td>
<td>742.00</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Floraia</td>
<td>1/15/2016</td>
<td>95.26</td>
<td>1873.00</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td>Napoli</td>
<td>1/27/2016</td>
<td>25.00</td>
<td>534.00</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Mariscal de Comandoni</td>
<td>1/27/2016</td>
<td>21.52</td>
<td>729.00</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>El Vergel</td>
<td>1/27/2016</td>
<td>22.52</td>
<td>609.00</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Cristobal Coles</td>
<td>1/29/2016</td>
<td>28.10</td>
<td>1009.00</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>El Poblado 1</td>
<td>1/29/2016</td>
<td>56.51</td>
<td>846.00</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Cali Mio</td>
<td>1/29/2016</td>
<td>20.00</td>
<td>3563.00</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Marroquín 3</td>
<td>1/29/2016</td>
<td>80.05</td>
<td>3563.00</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Puerto del Sol</td>
<td>1/29/2016</td>
<td>18.00</td>
<td>265.00</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Manuel</td>
<td>1/30/2016</td>
<td>23.50</td>
<td>549.00</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Molina</td>
<td>1/30/2016</td>
<td>73.18</td>
<td>1375.00</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>Vergel &amp; 3</td>
<td>4/10/2016</td>
<td>44.00</td>
<td>859.00</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Calipso and Yira Castro</td>
<td>4/10/2016</td>
<td>42.00</td>
<td>933.00</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>Narino</td>
<td>4/15/2016</td>
<td>28.00</td>
<td>1485.00</td>
<td>13</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 17 shows the overall spatial distribution of the rides. While most neighborhoods were covered only once, there were a few occasions where the neighborhood was repeated (Calipso), and in other situations second and even third rides bleed into the same study space. For the purpose of this paper, each ride was treated as a separate investigation and Calipso neighborhood was mapped to illustrate SVG (Figure 18).

discard dengue.” I1: “Then the cases that you all had were not registered?” CHW: “eh, no they were not registered. Unless, unless, for example they did the analysis and took the samples and did the analysis. normally the cases that came to us.”
Figure 17: SVG routes
Results

SVG Maps of risk factors in neighborhoods of interest

To demonstrate the SV and SVG approach, one neighborhood of particular interest is highlighted, Calipso. Calipso is a neighborhood of medium social strata (3), covering 1848 meters squared, 5455 total population, 30% Afro-Colombian, 47% male, received 4.3mm of rain during the study period, and contains a canal. While this neighborhood is characterized as having many of the social and physical vulnerabilities that should indicate a high dengue incidence, this is not evident in surveillance outcomes (chapters 2 and 3). Meeting with local community members, the community perception was positive.

To more fully investigate the situation in Calipso, the SVG for the neighborhood was transcribed and mapped using Wordmapper. From the resulting point shapefile, all words having some connection to water (canals, puddles, words describing flow, etc.) were queried and combined and a KDE performed. The KDE was again classified into ten equal bands (to approximate percentage risk) and these were contoured. The resulting risk map, therefore shows specific risk locations, and a more general risk surface based on water mentions where “risk” increases in 10% bands. The same querying

---

° “I1: [00:15:46] M: So, you have lived here since when? S1: Here in Calipso, it has been thirty years... And Calipso is a great neighborhood to live in. For everything, it is a very tranquil neighborhood. Well, like any neighborhood there is a lot of insecurities but it is very good to live here. I would recommend.”
could also have occurred for trash related terminology.

Figure 18 displays these risk contours along with the beige buffered SVG path. This path has been buffered to 50m, and can be used as a proxy for the denominator geography investigated in the neighborhood; anything of (visual) note occurring along this path would hopefully have been recorded in the comments. Spatial risk comments identified from the SVG are also displayed, again with a 50m buffer (in red). Seven labels (A to G) have been added so that comparisons can be made between the graphics. One key geographic feature, a drainage canal, is also identified and buffered to 100m. There is

Figure 18: mapping of the Calipso neighborhood; KDE of spatial risk classified into ten equal categories and contoured

\[ \text{Figure 18: mapping of the Calipso neighborhood; KDE of spatial risk classified into ten equal categories and contoured} \]

...
some overlap between the canal buffer and the spatial risk locations because specific features in and around the canal are identified, such as engineering work, bridges, and stagnant water sections all of which posed an elevated risk along the feature. The buffered canal, key spatial comments and the contoured “water” keywords overlap around areas A to C. A further common area between the key comments and contours is found at D. The remaining key mentions (E and G) are obviously not water specific.

![Diagram showing overlap between canal buffer and spatial risk locations](image)

*Figure 19: Google street view versus Spatial video: KDE of water classified into ten equal categories and contoured*

The SV was also used as a data source for the digitizing of risks, including standing water, trash accumulations, dogs, and other variables with potential connection to dengue, such as vegetation and evidence of poverty. Each risk location (for example standing water) was digitized as points to show the
spatial extent, while a weight was added to approximate the visual depth. While obviously open to user coding error, previous use of this method has proven successful, especially if the same coder is used which reduces the risk of inter-rater reliability. The same approach was also used for Google Street View (GSV) coverage of the area. This resulted 93 digitized water locations for SV and 99 for GSV, and 110 trash or tire locations for SV and 80 for GSV (Figure 19). Although GSV appears to be a single time frame, there were seven different periods of video collection within the coverage (7/2013, 10/2013, 1/2014, 9/2014, 1/2015, 12/2015, 4/2016).

Each digitized point was used as input into a KDE with a 50m kernel, weighted by the depth or volume of the risk. Each KDE was classified into ten equal categories and contoured (Figure 18 through Figure 20; the buffered SV path is still included for comparison purposes). In Figure 19, the best consistency for both the GSV and SV sources is for area E. While there is also some overlap with areas C and D, SV dominates in intensity. This pattern is also true for the general section of the map between B and D. While the area between A and B is proximate to the canal, there is not a lot of visible standing water outside of the channel. When comparing Figure 19 to Figure 20, areas A through D exhibit similarity between SVG and the two digitized sources of data, while interestingly the most intense visible section (E) for both GSV and SV is not captured by the SVG.
Figure 20: KDE of trash and tires classified into ten equal categories and contoured

There are similarities between the GSV and SV sources in terms of visible tires and trash, with the path between A and B being consistent. This is also true for the path between F and G, and the area to the southwest of D. Area C to D and especially E which had high levels of visible standing water only have lower levels of visible trash, with more being seen in these areas for SV). If this map is compared with Figure 20, there is consistency between spatial mentions involving trash along the path leading to B (from A) and the area around G.
Discussion

There is a rich tradition of spatial research into mosquito vectored disease, and especially dengue. Cali, Colombia is a large urban area with hypoendemic dengue which experienced outbreaks of Zika and chikungunya during the study period that has also benefited from several of these studies which have considered disease patterns, early web based surveillance systems, and the identification of fine scale environmental and social risks. Even so, in one of the most studied urban environments, there is still room for more granular investigation (both spatial and temporal), and the addition of context. In this paper, the various neighborhoods of Cali have been considered using a SVG approach, and for one neighborhood of interest the mapping of street level risks have been added using both SV and GSV as a data source.

SVG revealed many interesting aspects of the mosquito vectored disease landscape, ranging from spatially specific risks, such as a single drainage trench and the geographic variations along it, to more general observations about what was happening in Cali.

SVG can also be used to develop systematic risk mapping by identifying key locations of risk, or by analyzing compounds of similar risk describing words using a KDE. These maps were then compared to (water and trash) risks digitized from either GSV or SV. While there were some variations between the three sources, suggesting robustness in risk presence. For example, when combining all sources, the most likely areas of vulnerability include the path from A to B (water and trash), the path from C to D (water and trash), water risks around E, and trash between F and G.
86

Figure 21: Top 50% Contours of Risk by Disease. Surveillance data for Dengue, Chikungunya, and Zika are overlaid on the path buffer. Dengue October 2014 - April, 2016. Zika November 2015 - April, 2016. Chikungunya January 2015 - April, 2016.

While the value in using SV and SVG as a source to generate granular risk surfaces has been shown, the method can only be adequately judged if disease surveillance is overlaid. To do this, surveillance data for Dengue, Chikungunya, and Zika are overlaid on the path buffer (Figure 21). The same mapping procedure as for the previous risks is applied to these data (a KDE with a kernel of 100m, classified into ten equal categories and then contoured). Two immediate patterns emerge from Figure 21, the high intensity around E and then between C and D. It is also interesting to note that all three diseases exhibit markedly similar patterns, which is not surprising considering the same vector is involved with all three. If just this map based on disease cases had been made, these two areas would
certainly suggest priorities for intervention. However, there would be little insight as to what might be causing these “hotspots”, and therefore how to design the intervention or education strategy.

Returning to Figure 18 through Figure 20, point “E” generated the highest and most consistent (between GSV and SV) water related risk, with some visible trash (especially using SV). The path between C and D emerges as both spatial mentions of risk, and water related comments in Figure 19, and digitized water and trash risks for both GSV and SV sources, though SV dominates. Therefore, using these risk sources, it is suggested that potential causes for the disease hotspots are neighborhoods conditions resulting in visible standing water and trash accumulation.

An additional benefit of the SV and SVG approach is that the source material (transcribed comments and video) can be returned to for further validation of these insights. By returning to the video around E the main road is a commercial strip with several stores and activity on the street. The road also has small pocket parks with playsets (Figure 22, the location of this park is almost exactly where “E” is on the previous figures). Visible side roads (which were not driven) are densely packed with houses. Although there are sections of broken road with standing water puddles, most of the roads are well maintained. The quality of the buildings and road surface deteriorates more towards the canal. However, few of the dense residential streets were covered on the drive and it is here where the disease cases would be found.
Questions arise from the methods presented. Firstly, why were some of the highest risk areas identified through the three sources (A to B) not present in the disease risk surface even though this is a dense urban area? While there are possible explanations regarding a lack of surveillance data for people living in the hardest economic situations, and from the SVG mention is made that the “poverty” often increases closer to the canal edge, it could also be that the risks identified around the canal have an impact on a far wider area easily extending to C, D and E because of the flight range of the mosquito. Secondly, is Google Street View alone a suitable data source, thus negating the need for SV? While there is consistency in output, especially for the disease hotspot of E, the number of time periods included in GSV is an obvious concern. We have also compared how environments have changed in Cali between the GSV and SV dates, and for some neighborhoods there is considerable variation. Cali has more consistency in its visible poverty. Therefore, GSV is a reasonable source of insight if one acknowledges the variations in time frame as an uncontrolled confounding factor. However, for many disease-threatened areas GSV is simply not available, and this, along with being able to dictate when video data...
are collected, still suggest SV, whenever feasible, is the better source. Given the importance of E in terms of visible water risks, and disease occurrence, why was this section not described in detail during the SVG? Again, the benefit of the SV and SVG approach is that the source data can be returned to and examined for this exact stretch of the data collection. Returning to the spatial comments, the sole mention closest to E does in fact read, “Look at the stagnant water in the (channel).”

Limitations of the studies

We suspect systematic underreporting of by region according to access to health services related to SES as also reported by Sarti et al where in Colombia, incidence rates for confirmed dengue were 5.8 times higher in the independent study compared to local state and 3.5 times higher compared local levels. 217. Our sample consisted of an estimated 5% of total cases (MoH personal communication). Cases not able to be geocoded were excluded from analysis (%).

Environmental risks were collected in a cross-sectional manner and represent the time at which these were collected which may depend on: season (rainy versus dry), program funding, and political will. Interviews were open-ended and subject to personal biases and perceptions. Vector specialist, health care workers, and residents were recruited as guides and narrators of the videos, but the knowledge level varies by interview.

The safety and social acceptance of field work was questionable and the support of the local MoH and CHWs is greatly appreciated.

Conclusion

This paper has revealed the types of risks present at a geographic scale not commonly analyzed with regards to dengue or other mosquito vectored disease. The SV and SVG approach not only can be used to identify potential risks, but also, when overlaid with actual disease cases, be used as an ongoing data resource to see what was happening inside each hotspot, what it looks like, and how it is described.
This detail is invaluable when targeting limited intervention resources. For example, while there remain many unanswered questions for the Calipso neighborhood, including is there a surveillance “hole” closer to the canal, it might be prudent to target intervention into areas E and C. Given the insights revealed from the SV and SVG, improving drainage and road surface quality might have an immediate impact on disease reduction. However, as only the main (commercial) roads were driven, a return visit might include all residential streets to assess the conditions closer to where people live. People in the general vicinity should also be educated as to the real disease risk and encouraged to use building screens, mosquito repellant or at least wear clothing while outside.

The approach presented here should also be considered as only a first step. It provides a baseline that can be added to as seasons change, or at regular time periods to assess change, especially after any intervention. By adding further SVG, more depth can also be added to the neighborhood, and at the same time incorporate different perspectives. As such, this approach not only bridges the gap between researchers and practitioners, but it can also help involve community participation into the process. Community insight is not only invaluable in terms of explaining why risk occurs, by also at understanding local perception, beliefs and challenges that could affect education initiatives. Why do so few houses use screens? Why is there no local desire to control trash even around the home? Why do men do outside work without clothing even though mosquitoes are present? These are the types of questions that can be investigated through the SVG leading to more culturally appropriate intervention strategies.

Here the first field validation of geospatial narratives for vector-borne diseases in South America is presented. These findings are relevant to local epidemiology as these allow the local resources to be optimized for vector control. These methods can also be used in other areas to guide MOH in resource management.
Although the risk factors for vector-borne diseases are well described (poverty, trash, water accumulation, limited access to health care, vector abundance); however, if these risks are not measured and reported, it can be argued these will not improve. Local to Cali know that these risks exist in varying degrees across the city. Here these risks are measured and are associated with disease presence at the neighborhood level. These provide the basis to create community education and intervention plans and provide evidence to ask for additional support, promoting social justice and environment health.\textsuperscript{134,205}.
Chapter 5

Conclusions

Overarching Discussion with Conclusions and Limitations

Here the geographical, epidemiological and social aspects of dengue, chikungunya, and Zika outbreaks in Cali, Colombia from 2014-2016 are discussed. The strengths include the readily available ministry of health data, community participation, novel mapping methods, and environmental risk analysis. The limitations are passive reporting and areas of risk with low coverage due to underreporting and mistrust in the traditional health system. This analysis can be used to guide further control in the neighborhoods of risk as arboviral outbreaks are a continued threat in this increasingly globalized city. The results of the study are being socialized in the study area and to the community partners for continued feedback and future work.

Recommendations for Future Research

Future research should include serological testing in hotspot neighborhoods to find true incidence/prevalence. We hope to continue to strengthen ties between community leaders and MOH for vector control programs by neighborhood. Community members have expressed interest in fish vector control in la Laguna and community care of canals and discussion of logistics of conversion to soccer fields. Future mapping would focus on areas of research where changes are seen or there is an epidemiological study for serological testing.

We also are in the process of coding the risks for 8 of the more interesting neighborhood's and then compare the video to google street view to see how stable the risks have been over time and have
the spatial video risk coding for three other neighborhoods which could produce similar analyses. We continue to collect data with MOH and CHWs in Cali with four cameras and two microphones available and systems set up to send and receive data; translate and transcribe the interviews; and to create meta data and maps.

Most importantly, the summary of results will be distributed to collaborators and translated into future research foci.

Impact

These results should be used in conjunction with cost-based intervention estimates of disease burden averted to affect policy change. Creating an impact in a community after a research is done is one of the most essential aspects to be able to improve the health outcome. Compiling lists of recommended health outcomes, sources and determinant and using the results of their analysis to plan strategies provides Community health workers with tools and information to help their community. It is essential to define the responsibility of the different sectors of local services after a research is done to continue doing the work that is necessary. Due to the complexities of vector borne diseases, making sure that control efforts are adequately target within reasonable amount of time by primary health care systems is important.

A risks maps with the latest data can support the active search of new outbreaks, ensuring that the correct help and resources that I needed. Similarly, direct impact on policy and introduction of new technology is another great way of creating an impact to community health worker after a research is made.
Investments and overall coordinated global commitment are vital to coordinate activities for the bettering of the community, such as having a basic stock of medicines, vaccines and personal equipment can tackle an outbreak if needed. Where these researches are being conducted, it is important to communicate the information compiled to the community’s leaders for them to share. Providing the community with the drugs and vaccines needed is another one key activity to move forward. Another way to guarantee impact is planning a community forum or meeting were information and resources are given. Mobilizing the communities to get educated and coordinate advocacy material while ensuing harmony while collaborating with other partners and increasing financial support. Ensuring the data will be distributed using meetings, social media, websites, broadcasting and with simple to understand language can also make the impact greater. Getting feedback on how the situation was addressed can allow further knowledge on how to improve it. Enabling a good work environment for community health workers to ensure good productivity with manageable workload, supportive supervision, supplies and equipment, and respect from the community.
References


39. Branswell H, 2016. Experimental Zika vaccine is approved for clinical trials for first time in US. STAT: Reporting from the frontiers of health and medicine. Trend MD.
43. Steim M, Juri M, xed, a JD, Oria GI, Ramirez PG, 2013. Aechmea Distichantha (Bromeliaceae) epiphytes, potential new habitat for Aedes Aegypti and Culex Quinquefasciatus (Dipteria: Culicidae) collected in the province of Tucuman, Northwestern Argentina. The Florida Entomologist 96: 1202-1206.
60. Departamento Administrativo Nacional de Estadistica, 2005. [General Census].


219. U.S. Centers for Disease Control and Prevention, 2013. Community Health Assessment for Population Health Improvement: Resource of Most Frequently Recommended Health Outcomes and Determinants. Atlanta, GA.


Appendices

IRB approval letters

IRB Level I, category 2 approval for Protocol application #15-529 - please retain this email for your records

RAGS Research Compliance <researchcompliance@kent.edu>  Wed, Oct 7, 2015 at 8:47 AM
To: "James, Mark" <mjames22@kent.edu>, "akrysto1@kent.edu" <akrysto1@kent.edu>, "Curtis, Andrew" <acurtis3@kent.edu>

RE: Protocol #15-529 - entitled "Chikungunya and Dengue in Cali: Epidemiological and Geospatial Analyses"

We have assigned your application the following IRB number: 15-529. Please reference this number when corresponding with our office regarding your application.

The Kent State University Institutional Review Board has reviewed and approved your Application for Approval to Use Human Research Participants as Level I/Exempt from Annual review research. Your research project involves minimal risk to human subjects and meets the criteria for the following category of exemption under federal regulations:

- Exemption 2: Educational Tests, Surveys, Interviews, Public Behavior Observation

This application was approved on Oct 7, 2015 by the Kent State IRB. Please note that this approval only extends to Kent State Investigators (Mark James, Amy Krystosik, and Andrew Curtis). It is understood that the research will take place in Cali, Columbia and that the research has also been reviewed and approved by the ICESI University IRB.

***Submission of annual review reports is not required for Level I/Exempt projects. We do NOT stamp Level I protocol consent documents.

If any modifications are made in research design, methodology, or procedures that increase the risks to subjects or includes activities that do not fall within the approved exemption category, those modifications must be submitted to and approved by the IRB before implementation. Please contact an IRB discipline specific reviewer or the Office of Research Compliance to discuss the changes and whether a new application must be submitted. Visit our website for modification forms.

Kent State University has a Federal Wide Assurance on file with the Office for Human Research Protections (OHRC); FWA Number 00001853.

If you have any questions or concerns, please contact us at Researchcompliance@kent.edu or by phone at 330-672-2704 or 330.672.8058.

Respectfully,

Paulette Washko, MPH, CIP
Kent State University
Acta de Aprobación N° 061

Proyecto:

Chikungunya and Dengue in Cali: epidemiological and geospatial analyses

Sometido por: Diana María Dávalos, Mark A. James, Amy Robyn Krystosik, Robinson Pacheco, Sarita Rodríguez, Fernando Rosso.

El Comité de Ética de Investigación Humana de la Universidad Icesi, creado mediante la Resolución de Rectoría No. 763 del 13 de Abril del 2010, se rige por la Resolución 008430 del 04 de Octubre de 1993 del Ministerio de Salud de Colombia, por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud; la Resolución 2378 de 2008 del Ministerio de la Protección Social, por la cual se adoptan las Buenas Prácticas Clínicas para las instituciones que conducen investigación con medicamentos en seres humanos; los principios de la Asamblea Médica Mundial expuestos en su Declaración de Helsinki de 1964, última revisión en 2013; y el Código de Regulaciones Federales, título 45, parte 46, para la protección de sujetos humanos, del Departamento de Salud y Servicios Humanos de los Institutos Nacionales de Salud de los Estados Unidos 2000

Este Comité certifica que:

1. Sus miembros revisaron los siguientes documentos del presente proyecto:

   X Protocolo de Investigación

   Resumen del Proyecto

   Formato de consentimiento informado

   Folleto del investigador (si aplica)

   Resultados de evaluación por otros comités (si aplica)

   Instrumento de recolección de datos

   Carta de instrucciones a participantes

2. El presente proyecto fue evaluado y aprobado por el Comité:

3. Según las categorías de riesgo establecidas en el artículo 11 de la Resolución N° 008430 de 1993 del Ministerio de Salud, el presente estudio tiene la siguiente Clasificación de Riesgo:

   X Sin Riesgo

   Riesgo Mínimo

   Riesgo Mayor del Mínimo

4. Que las medidas que están siendo tomadas para proteger a los sujetos humanos son adecuadas.
5. La forma de obtener el consentimiento informado de los participantes en el estudio es adecuada. Según lo establecido en los artículos 15 y 16 de la Resolución 08430 de 1993. Con la descripción suministrada en la propuesta macro del proyecto se considera inicialmente que no requiere de un formato escrito para documentar el proceso de consentimiento informado, puesto que los métodos seleccionados para recolección de información permiten clasificar la investigación en la categoría de investigación sin riesgo. Sin embargo, los investigadores se han comprometido, y así lo enuncian en la propuesta revisada, a someter el protocolo detallado del estudio de cohortes que planean desarrollar, junto con las respectivas aprobaciones por los comités de las instituciones que sean seleccionadas para participar.

6. Este proyecto será revisado nuevamente en la próxima reunión plenaria del Comité, sin embargo, el Comité puede ser convocado a solicitud de algún miembro del Comité o se las directivas institucionales para revisar cualquier asunto relacionado con los derechos y el bienestar de los sujetos institucionales para revisar cualquier asunto relacionado con los derechos y el bienestar de los sujetos involucrados en este estudio.

7. Informará inmediatamente a las directivas institucionales:
   a. Todo desacato de los investigadores a las solicitudes del Comité.
   b. Cualquier suspensión o terminación de la aprobación por parte del Comité.

8. Informará inmediatamente a las directivas institucionales toda información que reciba acerca de:
   a. Lesiones a sujetos humanos.
      Problemas imprevistos que involucren riesgos para los sujetos u otras personas
   b. Cualquier cambio o modificación a este proyecto que haya sido revisado y aprobado por el Comité.

9. El presente proyecto ha sido aprobado por un periodo de 1 año a partir de la fecha de aprobación.
   Los proyectos de duración mayor a un año, deberán ser sometidos nuevamente con todos los documentos para revisión actualizados.

10. El investigador principal deberá informar al Comité
    a. Cualquier cambio que se proponga introducir en este proyecto. Estos cambios no podrá iniciarse sin la revisión y aprobación del Comité excepto cuando sean necesarios para eliminar peligros inminentes para los sujetos.
    b. Cualquier problema imprevisto que involucre riesgos para los sujetos u otros.
    c. Cualquier evento adverso serio dentro de las primeras 24 horas de ocurrido, al secretario (a) y al presidente.
    d. Cualquier conocimiento nuevo respecto al estudio, que pueda afectar la tasa riesgo/beneficio para los sujetos participantes
    e. Cualquier decisión tomada por otros comités de ética
f. La terminación prematura o suspensión del proyecto explicando la razón para esto
g. El investigador principal deberá presentar un informe al final del año de aprobación. Los proyectos de duración mayor a un año, deberán ser sometidos nuevamente con todos los documentos para revisión actualizados.

Firma: [Firma]
Fecha: 14 09 2015
Nombre: Yoseth Ariza-Araujo
Teléfono: 5552334 ext. 8140
Capacidad representativa: Presidente del Comité de Ética Humana
APPENDIX 1: INFORMED CONSENT FOR THE COMMUNITY HEALTH WORKER INTERVIEW

Informed consent form for men and women (> 18 years) participating in the project:
"CHIKUNGUNYA AND DENGUE IN CALI: Epidemiological and Geospatial ANALYSES."
As community storytellers in structured and recorded interviews.

<table>
<thead>
<tr>
<th>Location:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Principal Investigator: Mark A James.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Community health worker’s Name:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>DOB:</th>
<th>ID:</th>
<th>Sex:</th>
</tr>
</thead>
</table>

Introduction

Chikungunya is a disease that is in outbreak and is causing thousands of cases in this region, causing a deterioration in the overall health of different communities. Perhaps you or someone close to you has suffered from this disease, which is caused by a virus transmitted by the bite of a mosquito that has several common names, such as zancudo, mosquito, flechudo or weevil, among others. One aspect that can be investigated to ensure that fewer people become infected is to reduce risks such as mosquito bites.

EL INSTITUTO DE INVESTIGACIONES CLÍNICAS DE VALLE DE LILI AND KENT STATE UNIVERSITY AND LA RED DE SALUD LADERA are interested in the study of chikungunya. We plan to study and map the risks of transmitting the disease in your community. This study gives us information to choose a control measure to reduce the number of cases transmitted.

The duration of this study is 6 months. And prior to the execution of your duties, you will be trained by the epidemiological study staff.

Procedures and Functions
Your work in the study will 1) answer the questions in the structured and recorded interview and / or 2) be a neighborhood guide for mapping. For this you will be trained by qualified personnel prior to implementation of the protocol.

To record the interview and neighborhood, you must use the technique of "spatial-video software 'storytelling.'” This means you will narrate a video map of your community. After recording the video, you must help the staff to the code of points of interest and risks in the community to generate the map.

There is no risk to you beyond normal daily activities.

Risks and Benefits
If you decide to participate, you must keep in mind the following information:

a. The risk of infection by any vector transmitted disease, such as chikungunya or dengue, is the same as that presented to you in the community when areas of the body are exposed to the environment.
b. There exist no penalties if you refuse to participate in the study.
c. You can withdraw from the study when you find it convenient, and this decision will not cause you any type of penalty.
d. You must to be able to read and respond to a questionnaire related to this study on chikungunya to be eligible to participate in the study.
e. There will be no additional direct benefit to you but there could be general benefit to the society

Rights
If you decide to participate, you have the following rights:

a. The records which could identify you will be kept confidential according to the good clinical practices in research. Only the study staff, the institution or funding and regulatory agencies can access the data from this study, but they also must maintain confidentiality.
b. Your participation in this research is completely voluntary.
c. Any questions or concerns please communicate with Amy Robyn Krystosik (313 549 3826 or440 536 8523) who will inform you about your rights to participate or call the number of the local Ethics Committee, Comité Ética de La Red de Salud Ladera: 608 0124 PBX or the Kent State University IRB at (330-672-2704), where they will give you the information on the study.
With your signature, you certify that you have read or someone has read to you this informed consent form; that they have successfully resolved all your questions and that you voluntarily agree to participate in the study entitled: "CHIKUNGUNYA AND DENGUE IN CALI: Epidemiological, Geospatial, serotype and KAP ANALYSIS" as a community narrator in a structured and recorded interview.

<table>
<thead>
<tr>
<th>Community health worker’s name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID number:</td>
</tr>
<tr>
<td>Signature:</td>
</tr>
<tr>
<td>Place and date:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of the person who explained the study:</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID Number:</td>
</tr>
<tr>
<td>Signature:</td>
</tr>
<tr>
<td>Place and date:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of witness:</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID Number:</td>
</tr>
<tr>
<td>Signature:</td>
</tr>
<tr>
<td>Place and date:</td>
</tr>
</tbody>
</table>
Participant consent for audio and video recording

AUDIO / VIDEO GRABACIÓN
CONSENTIMIENTO INFORMADO
CHIKUNGUNYA Y DENGUE EN CALI: EPIDEMIOLÓGICA Y GEOESPACIAL ANÁLISIS
MARK A JAMES

Estoy de acuerdo en participar en una entrevista de audio/videos grabadas sobre "Chikungunya y dengue en Cali: epidemiológica y geoespacial análisis", como parte de este proyecto y para los fines de análisis de datos. Estoy de acuerdo en que Mark A James, Amy Robyn Krystosik, Robinson Pacheco, o Diana Dávalos puede grabar esta entrevista en forma de audio o video. La fecha, hora y lugar de la entrevista serán acordados mutuamente.

Nombres y apellidos: ____________________________ Firma: ____________________________ Cédula O ID: __________ Fecha: __________

Me han dicho que tengo el derecho de escuchar la grabación de la entrevista antes de su uso. He decidido que yo:

____ Sí deseo escuchar la grabación
____ NO deseo escuchar la grabación

Por favor firme ahora por debajo si no desea escuchar la grabación. Si desea escuchar la grabación, se le pedirá que firme después de escuchar a los mismos.

Mark A James / Amy Robyn Krystosik / Robinson Pacheco / Diana Dávalos puede / no puede ( marque uno) utilizar las grabaciones de audio / video hechas de mí. Copias o las originales se puede utilizar para:

____ este proyecto de investigación _____ presentación en reuniones profesionales _____ publicación

Nombres y apellidos: ____________________________ Firma: ____________________________ Cédula O ID: __________ Fecha: __________
Global Health Epidemiologist

Phone: 330-212-4719  E-mail: Akrysto1@kent.edu

SUMMARY: Epidemiologist with 8+ years of experience in international public health, epidemiology, tropical infectious disease surveillance, GIS, medical entomology and community outreach.

PROFESSIONAL EXPERIENCE

Epidemiologist: Chikungunya and Dengue in Cali, Colombia: epidemiological and geospatial analyses
Stanford University School of Medicine Dept. of Infectious Pediatrics- Stanford, CA  2016- Present
- Create four databases to manage, clean, and analyze field data collection
- Analyze preliminary data to verify lab results and present preliminary results
- Submitted 1 grant for continued research funding in Spatial Video Geonarratives in Urban Coastal Kenya

Epidemiology Principal Investigator: Chikungunya and Dengue in Cali, Colombia: epidemiological and geospatial analyses
Kent State University College of Public Health- Cleveland, OH  2012 - Present
- Mobilized local resources to improve vector control during chikungunya, dengue, and zika outbreaks in the field
- Collected more than 30,000 case data and local spatial data
- Targeted research in 2 foci through creation of local risk maps in ArcGIS
- Uncovered local trends in waste and vector management, useful for future infectious disease control programs

Epidemiology Researcher: Comparison of Efficacy of Attractive Toxic Sugar Bait Methods for Anopheles in Latin America
Caucaseco Scientific Research Center, Cali, Colombia  2014 – 2015
- Key research contributor for field testing of malaria control devices in Colombia
- Accelerated field work on local malaria vector species with advanced knowledge of novel laboratory techniques
- Enhanced presentation of 10+ IRB/laboratory/clinical protocols, peer reviewed articles, and marketing projects in collaboration with local health provider

Epidemiology Field Researcher
Tulane School of Public Health & Tropical Medicine- New Orleans, LA  2011 - 2012
- Advanced dengue control efforts through vector surveillance and data management of 30+ field sites
- Improved knowledge of health effects of Aedes as disease vectors utilizing parity dissections

Top 3 Finalist, “Be the Change: Save a Life” Public Health Challenge
May, 2011
ABC News and the Duke Global Health Institute
Project: Maba Tea
- Collaborated with colleagues on innovative solution to international healthcare problem
- Literature review on maternal health issues in India and solutions
Chagas’ Disease Eradication Researcher 2008 & 2010
Ohio University Tropical Disease Institute- Quito, Ecuador

- Designed and implemented local site program for outreach on vector-transmitted tropical parasitic disease, affecting local community of Loja, Ecuador
- Improved community health through epidemiology and entomology surveys, geographical information systems, and community education in 10+ rural communities with Ministry of Health
- Improved disease screening in rural schools and local hospitals through personal communication and public presentations to local clinics, schools, and households.

CERTIFICATIONS & SPECIALIZED EXPERTISE
- Peace Corps Community Health Outreach Volunteer, Matlala, Polokwane, South Africa, 2011 – 2012
- Global Health Immersion Course, KSU College of Public Health - Colombia, Panama, Guatemala
- Social & Behavioral Research and Conflict of Interest Training, CITI Collaborative Institutional Training Initiative 2014
- Creation and Validation of Risk Maps in the Context of Malaria Elimination, The Colombian Presidential Agency of International Cooperation and The Latin American Center of Malaria Investigation, 2014

KEY SKILLS
Epidemiological Experimental Design:
- Survey design: longitudinal, cross-sectional, epidemiological, entomological, KAP
- IRB protocol development
- Infectious disease vector assay development
Statistical Analysis and Biostatistics:
- Survival/mortality analysis, statistical modeling (logistic, ordinal, OLS, IV regressions), weighted means, frequencies, rates, trends, geographically weighted regressions
- Software: Stata, SAS, SPSS, ArcGIS, R
- Mapping and GIS systems
Project and Fieldwork Management:
- Project resource and data management
- On-site field staff training
- Collaboration with key stakeholders
Global Health and Development:
- Professional experience in Colombia, South Africa, Guatemala, Panama, Ecuador, El Salvador
- Infectious disease control and screening
- Tropical and infectious disease expertise: Dengue, Chagas’ Disease, Malaria, HIV/AIDS, Chikungunya, Zika
- International government collaborations
- Spanish (professional proficiency)
Scientific Communication and Public Education:
- Scientific writing and literature reviews
- Grant writing
- Educational content creation and outreach
- International teaching

PROFESSIONAL MEMBERSHIPS
- EpiCore
- American Society of Tropical Medicine and Hygiene (ACME & ACGH)
- Public Health Student Alliance, Kent State University College of Public Health

EDUCATION
Doctor of Philosophy of Public Health: Epidemiology December, 2016
Kent State University College of Public Health- Cleveland Ohio

Master of Public Health: International Health and Development December, 2011
Tulane University School of Public Health and Tropical Medicine- New Orleans, Louisiana

Bachelor of Science in Chemistry: Biochemistry May, 2010
John Carroll University- Cleveland, Ohio