ENVIRONMENTAL RISKS OF CHAGAS DISEASE IN THE RIO GRANDE VALLEY, TEXAS

A Thesis

by

BARA SAFAROVA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF URBAN PLANNING

Chair of Committee,	Cecilia H. Giusti
Committee Members,	Shannon S. VanZandt
	Stephen M. Caffey
Head of Department,	Shannon S. Van Zandt

May 2019

Major Subject: Urban and Regional Planning

ABSTRACT

Chagas disease is endemic in Latin America, where environmental risk predictors have been studied to inform an effective vector control strategy to limit its spread. In recent years, an increase in the prevalence of Chagas disease in canine, which presents a major risk to human health, was identified in the southern United States. However, in the U.S. little is known about environmental risks that could be used for vector control of Kissing bugs, which carry the parasite of the disease (*Trypanosoma cruzi)*. In this study, I compiled a spatial database of secondary data to identify environmental risks associated with the prevalence of Chagas disease in canines in the Rio Grande Valley in South Texas. Locations of 100 lots from a pilot study that determined a 19.6% prevalence level of *T. cruzi* among 209 canines, were used to collect environmental, land development, and socio-economic secondary data surrounding each lot. Results of a logistical regression showed the following increase in the odds of a positive lot: nearly three times increase for lots within a 1-mile of a natural area (*p-value* 0.089), presence of a half, or one, adjacent unconstructed lot increased the odds nearly five times (*p-value* 0.020), and the presence of more than one canine in a lot increased the odds more than three times (*p-values* 0.078, 0.98, and 0.001 for two, three, and four canines per lot respectively). The results corroborate previous studies that were conducted in Mexico, but further research is needed to understand the role of disinvestment in rendering environmental, land development, and canine characteristics as a risk of Chagas disease. Because no vaccine exists, planning policy recommendations based on vector control were formulated to prevent the spread of the disease: growth boundaries to limit residential construction in proximity to natural areas, ordinances requiring maintenance of vacant and abandoned lots in peri-urban areas, and a policy limiting the number of canines per lot in peri-urban areas.

DEDICATION

To my parents Zlaťka Šafářová and Petr Šafář, who gave me the freedom and the means to pursue knowledge under my own direction. Your trust in my abilities has been an underlying source of confidence in my pursuits. Děkuji!

ACKNOWLEDGEMENTS

I would like to thank the whole One Health team for the expertise and resources they shared with me during the research project and beyond. I would like to thank Ester Carbajal and Italo Zecca for their help, local and expert knowledge, and dedication during the fieldwork phase of this project – I could not have done this without you. I also would not have been able to spend enough time in the field without the generous support and warm welcome of Dr. Ann Millard and Dr. Isidore Flores, who opened their doors for me and helped me navigate both the area and the research. I would like to thank Dr. Rachel Curtis-Robles for her superb citizen outreach campaign that led to my introduction to the world of vector-borne disease, the Kissing bug, and the research following my initial encounter with the vector. From the One Health team, last but not least, I would like to express my gratitude to Dr. Sarah Hamer and Dr. Gabriel Hamer for their research input, patience, advice, and resources. I would also like to thank my colleagues at the Landscape Architecture and Urban Planning department at Texas A&M University for their guidance in navigating statistical analyses: Maria Perez, Alexander Abuabara, and Dr. Nathanael Rosenheim – thank you all!

The thesis would not have materialized without my thesis chair, Dr. Cecilia Giusti, who entrusted me with this project and invested resources, time, and knowledge in guiding me from the beginning, through many dead-ends and re-starts, until the end of the process. Dr. Stephen Caffey helped me see the big picture, and Dr. Shannon Van Zandt helped me navigate the methods and planning implications, although mistakes are all my own.

Last, I want to thank my partner Tatum Lau. I am grateful for your generosity with sharing your research expertise on green infrastructure, which was a timely gift that shaped my understanding of the project. And second, I'd like to thank you for your patience, support, and help with the biggest and the smallest things.

CONTRIBUTORS AND FUNDING SOURCES

Faculty committee recognition

This thesis was supervised by a thesis committee consisting of Dr. Cecilia Giusti [advisor] and Dr. Shannon Van Zandt of the Department of Landscape Architecture and Urban Planning and Dr. Stephen Caffey of the Department of Architecture.

Student/collaborator contributions

The work in this thesis was part of a larger research project (A multidisciplinary epidemiological approach to mitigate the human and animal health burden of Chagas disease across a transnational gradient), with which several Texas A&M departments were associated: Biology, Entomology, Geography, Landscape Architecture and Urban Planning, Veterinary Integrative Biosciences and Texas A&M Health Science Center department of Health Promotion and Community Health Sciences.

The data analyzed in this thesis were provided by the Principal Investigator of the larger study, Dr. Sarah Hamer. The data are a result of analyses conducted in by a team

of authors: Rachel Curtis-Robles, Italo B. Zecca, Valery Roman-Cruz of the College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas; Ester S. Carbajal of the Department of Entomology, Texas A&M University, College Station, Texas and the International Valley Health Institute, Edinburg, Texas; Lisa D. Auckland of the College of Veterinary Medicine and Biomedical Sciences, Texas A&M University; Isidore Flores of the Department of Entomology, Texas A&M University, College Station, Texas, Ann V. Millard of the School of Public Health, Texas A&M Health Science Center, McAllen, Texas; and Sarah A. Hamer of the College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas; and were published in (2017). Maria Perez, a doctoral student in the Department of Landscape Architecture and Urban Planning at Texas A&M University in College Station, Texas provided assistance and guidance with the statistical analysis, however any mistakes are the sole responsibility of the author. All other work conducted for the thesis was completed by the student independently.

Funding sources

This study was partially funded by an internal Texas A&M University One Health Grand Challenge Seed Grant; and in part by an internal Texas A&M University College of

Architecture Diversity Council Teaching Assistantship. The contents of the thesis are solely the responsibility of the author and do not necessarily represent the official views of the Texas A&M One Health Initiative, or the Texas A&M College of Architecture Diversity Council.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1 INTRODUCTION

Chagas disease is a vector-borne[1] disease transmitted by the "Kissing bug" insect, with 11 known species in the United States, for example *Triatoma recurva*, *T. rubida*, and *T. gerstaeckeri,* among others. The disease is widespread and well documented throughout Latin and Central Americas and increasingly documented in southern United States (Bern et al., 2011). Peri-domestic transmission cycle between canines and the vector has been confirmed in Texas (Kjos et al., 2013; Kjos et al., 2009). South Texas was determined as the area with highest risk of Chagas disease based on analyses of environmental characteristics (Sarkar et al., 2010). Even though direct transmission from canines to humans is unlikely, infected canines present a public health risk for transmission to people because they can serve as reservoirs hosts (Estrada-Franco et al., 2006; Ramírez et al., 2013a). A recent study of peri-urban, unincorporated settlements in South Texas showed a higher prevalence of Chagas disease in canines and anticipated that the contributing factors were the exposure to sylvatic environments, free-roaming

[1] "Vectorborne diseases, such as malaria, are those in which an organism, typically insects, ticks, or mites, carry a pathogen from one host to another" (NIEHS, 2017, para 1).

canines, and substandard housing (Curtis-Robles et al., 2017). The current study explores which environmental and land development characteristics explain the increase in the prevalence of Chagas disease in canines in South Texas.

1.1 Purpose and Objectives

The purpose of the study is to identify environmental and land development risks of Chagas disease and formulate planning policy recommendations to reduce the risk of the disease in South Texas, which may be applicable across similar conditions in the Southwest. To do this, three objectives were formulated. The first objective is to spatialize the prevalence of Chagas disease in canines, based on the results of a recent pilot study conducted as part of the One Health Initiative at Texas A&M University. The mapping of Chagas disease in canines to specific residential addresses will provide the basis for the assessment of the natural and built environment surrounding each sampled case. The second objective is to examine the relationship between the prevalence of the disease and both the environmental and land development characteristics. A statistical analysis of data, such as land cover types, vacancy rates, and lot sizes will illustrate the relationship between the factors and the disease prevalence. The third objective is to formulate planning policy recommendations based on the results of the statistical analysis to limit the conditions associated with higher prevalence of Chagas disease in canines.

The study was conducted as part of the One Health Initiative at Texas A&M University. The interdisciplinary team consisted of researchers from Texas A&M departments of Biology, College of Agriculture – Entomology, College of Geosciences – Geography, College of Architecture – Landscape Architecture and Urban Planning, School of Rural Public Health – Health Promotion & Community Health, College of Veterinary Medicine – Veterinary Integrative Biosciences. The study was funded by One Health Initiative seed grant.

1.2 Study Significance

This thesis contributes to the field of urban planning and public health in three ways. First, the study applies the methods and variables of a regional-scale spatial analyses of environmental risk characteristics of Chagas disease conducted in Mexico, to the South Texas context. Secondly, the study contributes to the growing literature on health benefits of access to green open spaces by contrasting benefits with possible epidemiological threats in peri-urban, low-income locations. Third, the study offers practical application for vector control through planning policy. Policy recommendations could be applicable in other regions where poverty overlaps with the demand for peri-urban residential developments, and where Chagas disease risk has been identified.

1.3 Area of Study

The study is located in South Texas in the Rio Grande Valley (RGV) where Chagas disease has been confirmed by previous studies and where other environmental, socio-economic, and land development conditions increase the risk of transmission. The RGV region is located in the easternmost part of the U.S. – Mexico border. The risk of vector-borne disease in the U.S. – Mexico trans-border region increases due to the migration and transportation of wildlife, humans, and goods (Esteve-Gassent et al., 2014), and is expected to increase over time with climate change and associated hazards (Garza et al., 2014).

The region has historically had a Latinx population majority (Garza, 2003), which has been steadily increasing to over 88% of total local population in 2014 (Pew Research Center, 2016). The RGV is one of the fastest growing regions in the U.S. (RGV grew 75% to 1.22 million, in contrast U.S. grew 24% during 1990 – 2009) and also one of the poorest: per capita income was about $13,500 in 2010, approximately half of the national level (Ryabov & Merino, 2017).

The RGV comprises of four counties: Cameron, Hidalgo, Starr, and Willacy (Fig. 1), where several peri-urban residential land development strategies emerged over time to accommodate the low-income population: mainly the so called colonias and Model Subdivisions. Colonias are small settlements, typically located outside city limits in county land at various stages of development, that historically have lacked potable water, electricity, public services and have had substandard housing conditions (Davies & Holz, 1992; Donelson & Esparza, 2016; Ward, 1999). Of the total of 2,019 colonias identified along the U.S. – Mexico border in 2006, 90% were located in Texas, with highest concentration in three of the four RGV counties: Cameron, Hidalgo, and Starr (The Colonia Initiatives Program, 2006). Model Subdivisions are characterized by similar or worse housing conditions than in the colonias (Durst & Ward, 2016), and are located further from city centers, on larger lots (Durst, 2016). Hidalgo county contains the highest concentration of Model Subdivisions (Durst, 2016).

Overall, these socio-economic, environmental, and land development characteristics render peri-urban settlements in South Texas vulnerable to health risks (Bogolasky & Ward, 2018), and epidemiological threats such as Chagas disease (Curtis-Robles et al., 2017).

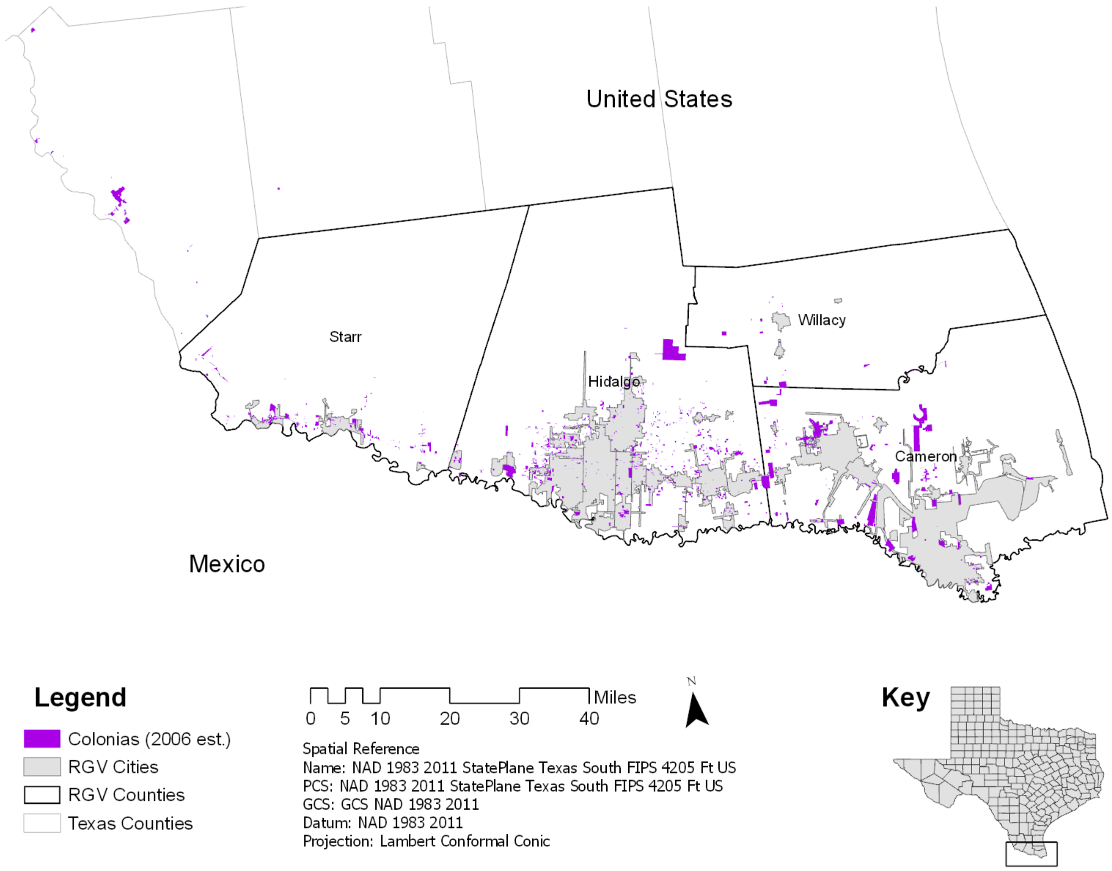

Figure 1: Area of study.
Data for RGV Cities, RGV Counties, and Texas Counties from U.S. Census Bureau (2016), and for Colonias from U.S, Geological Survey (Parcher & Humberson, 2007).

1.4 Structure of the Thesis

Section 2 reviews the literature and prior studies on the relationship between Chagas disease and the surrounding environment, including a review of peri-urban residential developments that may pose risks for vector-borne disease transmission in South Texas. Section 3 describes the operationalization and data collection pertaining to risk characteristics and the statistical method of assessment. Section 4 lists the results of the descriptive and statistical analyses. The descriptive analysis describes the distribution of Chagas disease with regards to various land cover and land development characteristics. The statistical analysis outlines results of logistical regression. Section 5 discusses the results with respect to prior studies and offers recommendations for planning policy and suggestions for future research.

2 LITERATURE REVIEW

In the first part, the literature review frames the study within the research on the relationship between health and green open spaces. The section is followed by an overview of Chagas disease, its burden, transmission cycle, and a review of prior research on the role of the environment in the spread of Chagas disease in the United States and in Mexico. The second section discusses types of residential developments in the U.S. and in particular in South Texas that could present a risk for Chagas disease transmission. The literature review concludes with a summary of potential risks that should be assessed by future research on Chagas disease in South Texas and elsewhere. Characteristics that are in the purview of urban planning policy, and thus could serve as a basis for planning policy recommendations for vector control, are highlighted.

2.1 Health and Green Open Spaces

Human settlements are increasingly more exposed to green open spaces in the United States, as housing developments are increasingly fragmented and built further from city centers (Durst, 2016; Durst & Wegmann, 2017; Ward & Peters, 2007). The planning literature views access to green open spaces as predominantly beneficial to health and well-being: increase in healing, self-reported well-being, reduction in obesity (Buehler, 2010; Cosco et al., 2014), improved mental health, and reduced mortality from cardio-vascular disease (Hartig et al., 2014; Shanahan et al., 2015; World Health Organization, 2016). However, the majority of studies have been conducted in high income countries (Markevych et al., 2017), fewer in the Developing countries, or low-income regions (cities, neighborhoods) suffering from disinvestment (Krellenberg et al., 2014; Shih, 2017; Xu et al., 2016). When combined with poverty, proximity to green open spaces holds risks associated with human exposure to wildlife and pathogens such as zoonic[2] and vector-borne diseases (Hartig et al., 2014). The major health risk is epidemiological as some green open spaces serves as a reservoir of wildlife and diseases, which can be transmitted to domestic animals or humans (World Health Organization, 2016).

[2] "Zoonotic Diseases (also known as zoonoses) are caused by infections that are shared between animals and people" (CDC, 2018, para 1).

2.2 Chagas Disease

Chagas is a vector-borne, zoonotic disease endemic in Central and South America, where approximately 8 million people are estimated to be infected (CDC, 2013), of which 20-30% will develop life-threatening conditions (Bern et al., 2011). The disease is also considered endemic in southern United States, particularly in Texas (Sarkar et al., 2010). The parasite causing the disease, *Trypanosoma cruzi*, has been identified in Kissing bugs (Fig. 2), and animal and human hosts in Texas since 1930s (Burkholder et al., 1980; Packchanian, 1939; Packchanian, 1942; Williams et al., 1977). Recent influx in studies of Chagas disease calls for more attention to the disease, as it is becoming a health risk to humans (Barr, 2009; Curtis-Robles et al., 2017; Tenney et al., 2014; Wozniak et al., 2015). Human contraction of Chagas disease in the U.S. has been attributed to prior residence in Latin America (CDC, 2013), however locally transmitted human cases have been identified (Cantey et al., 2012; Garcia et al., 2015; Garcia et al., 2016; Gunter et al., 2017).

Figure 2: Photo of an adult male *Triatoma gerstaeckeri* (species of a Kissing Bug) and 1st instar nymph, next to a penny.
Photographed by Gabriel L. Hamer. Reprinted with permission.

2.2.1 Chagas Transmission Cycle

T. cruzi is carried by a Kissing bug and transmitted through the contact between the feces of the bug after feeding and a bite site onto a range of hosts in the sylvatic cycle such as raccoons, woodrat and opossum (Bern et al., 2011; Hodo & Hamer, 2017; WHO Expert Committee, 2002). Kissing bugs colonize nests of wild animals and target animals kept outdoors in proximity to dwellings, such as canines and non-human primates (Hodo & Hamer, 2017). Active peri-domestic transmission cycles between Kissing bugs and domestic canines have been established in southern U.S. (Beard et al., 2003; Kjos et al., 2009; Sarkar et al., 2010). Infected canines increase the risk of human infection because they serve as reservoir hosts (Gürtler & Cardinal, 2015; Raghavan et al., 2015; Ramírez et al., 2013b). Kissing bugs have been found to invade dwellings in south Texas (Wozniak et al., 2015). Where housing expands into sylvatic environments, the risk of housing infestation, and of transmission to humans, are increased (Bern et al., 2011). Efforts to reduce the risk of Chagas disease must be based on vector control because no vaccine exists (Sarkar et al., 2010). Vector control preventative measures are also the most cost-effective means of combatting the spread of the disease (Hanford et al., 2007).

2.3 Prior Studies: Risk Characteristics

Environmental risks associated with the transmission of Chagas disease from vectors to peri-domestic and domestic locations have been studied primarily in Latin

America, where the disease is endemic. The spread of the disease, and the vector, depend on local socio-cultural and ecological conditions (Enger et al., 2004). The overwhelming factor for Tropical Neglected diseases, such as the Chagas, is low socio-economic status (Hotez et al., 2012).

Only one study, known to the author, assessed environmental risks of Chagas disease in canines in the United States at the regional scale. Raghavan et al. (2015) analyzed landcover, land use, and housing variables, and found the following to be significantly associated with Chagas prevalence in canines: housing built prior to 1980; rural location; and the total number of owner occupied housing units where householder was Hispanic or Latino. The geographic distribution of the cases was predominantly in central and northeastern part of Texas. At a domestic scale, Klotz et al. (2016) reported on the housing conditions of 10 cases of Kissing bug intrusions into homes and bites of humans, in Southwestern United States. These occurred usually in older homes without modern sealing and screening, but a small sample size prevented a statistical analysis. At a large ecological scale, Sarkar et al. (2010) analyzed data from cases in Mexico and the United States to estimate suitable habitat for Kissing bugs. They overlaid bioclimatic characteristics such as temperature, precipitation, and topographical factors with Kissing bug occurrences to establish the ecological risk of the disease in North America. They found that the risk was highest in South Texas. Similarly, Garza et al. (2014) used bioclimatic variables (temperature, precipitation, etc.) to predict suitable habitat for the vector across continental Mexico and U.S. under several future scenarios of climate change. They found that the potential distribution of suitable habitat for the Kissing bug shifts towards northern and eastern regions in Mexico and the United States.

Because only a few studies of environmental risks of Chagas disease have been conducted in the U.S., and only one at an appropriate regional scale for this current study, regional-scale studies located in Mexico were reviewed to identify risks that may be relevant in South Texas.

Ramirez-Sierra et al. (2010) found that housing infestation with Kissing bugs progressed from peripheral locations towards the center in four villages in the Yucatan peninsula in Mexico. The study relied on the collection of vectors by local inhabitants over the period of two years (2006 – 2008). The collected bugs were geocoded to residential addresses and distances to the edges of villages were measured as distances to the surrounding bushes based on satellite images and observations. Using statistical analyses, the authors determined that the density of the vector was more than double within the peripheral zone (less than 80m from the edge) than in the more central zones of villages. These results suggest that residential developed areas protect from vector-borne disease such as Chagas and that development patterns fragmented into smaller settlements may pose a risk to health. Barbu et al. (2010), together with two researchers from the previous study, using the same data from one of the villages in the Yucatan peninsula, used a selection model approach to explain dispersal dynamics of the vector. They found that over half (55%) of the infesting vectors originated in the peri-domestic areas within the village, and the rest arrived from the sylvatic environment beyond the edge of the settlement. They also found that the bugs were 5 – 15 times more likely to disperse towards houses than to peri-domestic spaces.

Dumonteil et al. (2013), together with some authors of the previous studies executed a follow up study two years later (2010 -2011) in three of the four villages in the Yucatan peninsula. They used random sampling of households, survey, and community collection of vectors. Infestation was defined as at least one vector found inside a house. Through the survey, they collected characteristics of housing, peri-domestic spaces, cultural, and socio-economic status. Using logistical regression analysis, they determined characteristics that best explain housing infestation with Kissing bugs. Their findings confirmed that peripheral location significantly increased the risk of infestation, but also found that keeping chickens in coops, cleaning trash from the peri-domestic area, and keeping more than two canines increased the risk of infestation. Risk factors of lesser importance were the proximity to street lights, and the presence of rock piles in the peri-domestic space.

Pacheco-Tucuch et al. (2012) using data collected in the previous studies determined that Kissing bugs were attracted to public street lights in the three villages in the Yucatan peninsula in Mexico. Housing closer to street lights was 1.64 times more likely to be infested (having at least one Kissing bug within the house) than housing located further from street lights in all three villages.

Ramsey et al. (2005) assessed environmental factors associated with housing infestation by Kissing bugs in the metropolitan area of Cuernavaca, Mexico. They stratified data collection – conducted by research personnel – by socio-economic status of population catchment units, which were also distributed by altitude (altitude of homes was associated with wealth). They found that poorer areas had double the rates of infestation than wealthier areas. In a multivariate regression analysis they found that the most significant predictor of infestation were: lower socio-economic status/lower altitude; large lots (garden larger than 80sqm); dogs able to enter and exit houses; occurrence of wild and domesticated animals (squirrel, opossum, pig); and at least one adjacent unconstructed lot.

Enger et al. (2004) collected household and environmental characteristics, and collected vectors across the town of Chalcatzingo, Morelos, Mexico. They determined levels of housing, peridomestic, and overall infestation with the vectors and used logistic regression to identify variables that explain each type of infestation. Domestic infestation was explained by the presence of agricultural products, junk piles, domestic animals in yard (rabbits), and the lack of bed nets. Peridomestic infestation was significantly associated with the presence of junk piles, and the number of dogs, cats, and rabbits in yard. Overall levels of infestation were associated with the presence of junk piles, agricultural products, and the number of domestic animals in yard (cats, rabbits, and fowl).

At a large, bio-climatic scale, researchers identified characteristics pertaining to climatic conditions that are associated with the presence of Kissing bugs. Ramsey et al. (2000) relied on community collection of the vector and trained personnel surveys (1996 – 1998) across municipalities within the state of Oaxaca, Mexico to determine the distribution of domestic Kissing bugs, and Chagas disease transmission. They found that different species of the vector were found in mostly separate bio-climatic regions.

Dumonteil et al. (2002) studied housing and bio-climatic characteristics across 23 villages dispersed across the Yucatan peninsula in Mexico during the year 1999 –

2000. They found that the type of housing had no statistically significant relationship with housing infestation, however bug abundance was significantly associated with the types of vegetation that was typical for northern locations.

Several scholars have used ecological niche modelling to assess large scale regional climatic and environmental conditions such as temperature, rainfall, etc. Several studies were conducted in Mexico (Costa & Peterson, 2012; Ramsey et al., 2015) and elsewhere (Peterson Townsend et al., 2011). Other researchers employed mathematical models to assess the prevalence of Chagas disease (Nouvellet et al., 2015). These studies suggest that climatic and large-scale environmental conditions undergird the presence of the vector and the disease.

The objective of the current study was to assess the environmental risk characteristics that may be associated with the spread of Chagas disease in peri-urban locations in South Texas. Because no study has been conducted thus far in this geographic context, the following review of peri-urban settlements was undertaken to translate risk characteristics identified in studies in Mexico to the South Texas context.

2.4 Peri-urban Developments in South Texas at Risk

In addition to dispersed homesteads and ranches within the extraterritorial jurisdiction of county land, two types of peri-urban settlements emerged in South Texas: unincorporated colonias and the so called Model Subdivisions. Additionally Informal Homestead Subdivisions were identified in the extraterritorial jurisdiction of metropolitan areas in the interior of Texas (Ward & Peters, 2007). The following review is limited to South Texas peri-urban settlements, where this study is located.

2.4.1 Colonias at Risk

Colonias are settlements which proliferated along the U.S. southern border predominantly in Texas since the 1960s and 1970s due to the lack of affordable housing options for a growing low-income population (Davies & Holz, 1992). Colonias were, from the onset, characterized by substandard housing built through self-managed, or self-help construction; a lack of, or a sporadic provision of electricity, water, and sewage infrastructure (Ward, 1999). These conditions made colonias vulnerable to health risks (Bogolasky & Ward, 2018). Over time, through incremental housing improvements, and with federal and state funding for the installation of infrastructure, some colonias' conditions improved (Durst & Ward, 2014; Giusti & Estevez, 2011). However, several characteristics increase the vulnerability of colonias to vector-borne diseases.

Typically, colonias are smaller in size than cities, and dispersed beyond the city limits – fragmented from the developed land of larger cities, which leaves them more exposed to green open spaces, wildlife, and pathogens. A combination of regulatory and market conditions further increases the risk of vector borne diseases. Ward and Carew (2000) identified high rates of absentee owners in colonias in the late 1990s. Lots with absentee owners had vacant buildings or were unconstructed all together. Similarly, more recent findings of an increase in renting (Durst, 2014b), abandonment, and vacancies (Durst & Ward, 2015). Both of these conditions likely render the lots more attractive for wildlife and vectors. Because municipalities have been underbinding colonias to avoid incorporating them within the city limits (Durst, 2014a), many colonias are excluded from municipal maintenance of public land, including trash collection, resulting in overgrown right-of-way areas and easements. Moreover, the underbinding also means that colonias land owners (including owners of vacant and abandoned lots) are excluded from public health municipal ordinances that would require them to remove rubbish, brush, lumber, and mow weeds and grasses, such as for example *Sec 54-62.*

Rubbish; weeds ordinance effective in the city of Brownsville in Cameron county: Code 1971, § 16-32; Ord. No. 93-562-C, § 1, 8-31-1993 (Brownsville, 1993).

2.4.2 Model Subdivisions at Risk

In 1989, a Model Subdivision (MS) rule was passed to curb the development of further colonias, requiring developers to install basic physical infrastructure prior to selling lots (Durst, 2016). The Model Subdivision rules also require the transfer of ownership of public streets, squares, and easements (among others) to a municipality. This means that right-of-way areas will be maintained by a municipality, reducing the opportunity for wildlife to establish. On the other hand, MSs are platted into larger lots, which was identified as a risk characteristic. Furthermore, because Model Subdivisions are newer, some may have higher rates of unconstructed lots which was shown to increase the risk of Chagas disease. Housing conditions in MSs resemble those in earlier colonias, though some are worse, because occupants have less financial resources left for housing construction due to higher costs of serviced lots (Durst & Ward, 2016). Additionally, majority of MSs (95%) are located in unincroporated areas (whereas only 65% of colonias), further from city limits, and have been platted into larger lots (Durst, 2016). These characteristics increase the exposure to sylvatic environment and access of the Kissing bug to peri-domestic, even domestic environments, and thus overall increase the risk of Chagas disease transmission. On the other hand, the increase in the upfront costs of infrastructure improvements in MSs, has meant that developers tend to undertake larger subdivisions (Durst, 2016). Larger subdivisions leave less properties exposed to the surrounding agricultural or natural land, which should serve to reduce the risk of transmission of Chagas to the interior of the settlements. In other words, because Kissing bugs spread into settlements from the periphery (Ramirez-Sierra et al., 2010), larger settlements such as MSs should serve as protectors from the risks of vector-borne diseases such as Chagas disease.

2.4.3 Peri-urban Developments in the U.S.

Peri-urban settlements can be found across the U.S., particularly in the South, where Chagas disease is endemic and Kissing bugs have been found. For example, peri-urban settlements and a range of informal settlements were identified in the interior of Texas (Durst & Wegmann, 2017), North Carolina (Ward & Peters, 2007), California and Ohio (Anderson, 2007), so called Wildcat subdivisions were found in Arizona (Christensen et al., 2006); other types of unincorporated settlements were found in the delta of Mississippi (Aiken, 1987); across the South (Lichter et al., 2007); in the West and Northwest of the U.S. (Shultz & Groy, 1988). Some of these states have also seen the presence of Triatomine vector, or vectors were reported in houses, and in ten southern states vectors were infected with *T. cruzi* (shown cross-hatched, horizontally hatched, and in green fill, respectively in Fig. 3. The overlap of peri-urban developments and occurrences of Kissing bug vector across the southern states of the U.S. suggests that more research may be needed to address the possible risks to public health associated with the fragmented and unincorporated types of development, particularly in poverty-burdened regions.

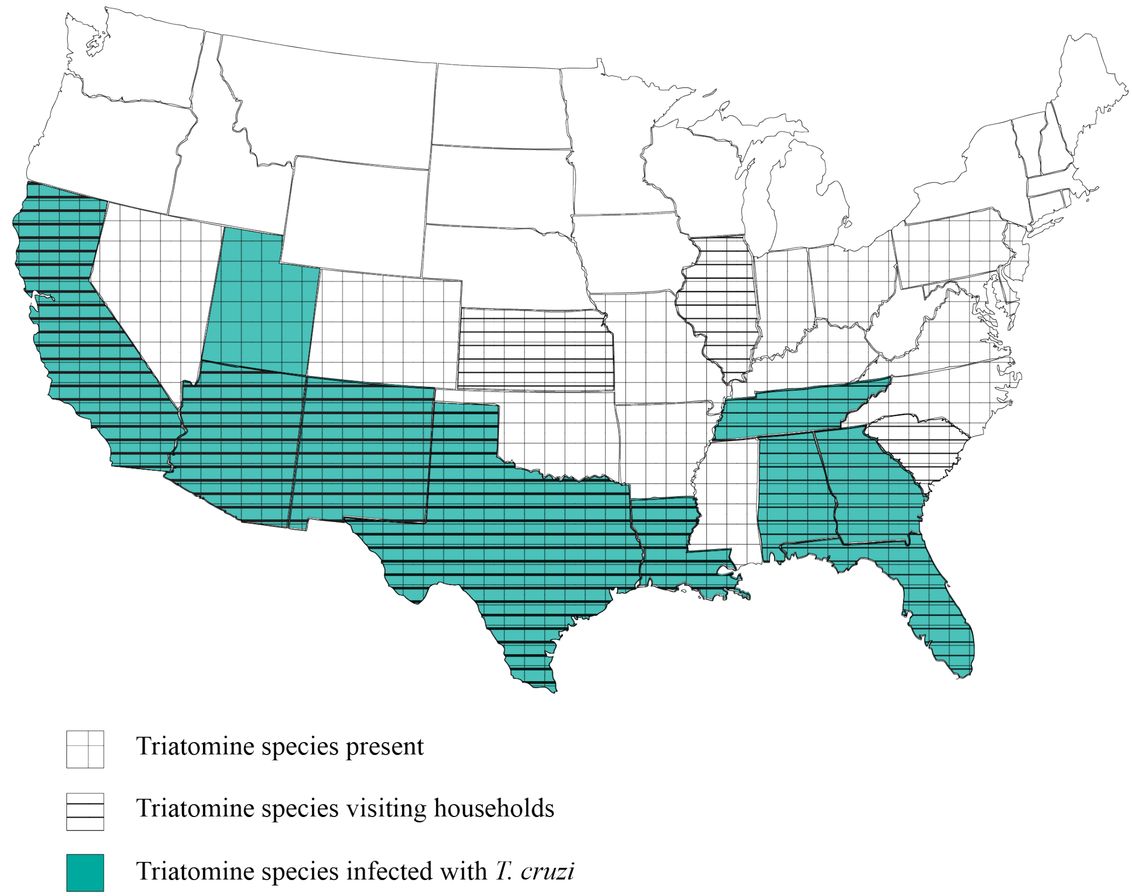

Figure 3: Geographic distribution of triatomine vectors, vectors visiting households, and vectors infected with *T. cruzi* across the United States.
Adapted using data for state boundaries from U.S. Census Bureau (2016), and data for Triatomine species distributions from Zeledon et al. (2012).

2.5 Chagas Disease Environmental Risks in South Texas

In the absence of prior studies of environmental risks associated with Chagas disease in South Texas, this section synthesized research conducted in Mexico with research on peri-urban settlements in South Texas to identify risks to be assessed by the current and future studies.

The location of the current study falls within the high risk region identified by large-scale ecological models: South Texas rural areas with Latinx-majority population. Risk characteristics identified in Mexico fall into several broad groups. The first group pertains to land development processes and exposure to green open spaces such as: peripheral location (Dumonteil et al., 2013; Ramirez-Sierra et al., 2010); peri-domestic areas and surrounding green open spaces (Barbu et al., 2010); adjacent unconstructed lots, and large lots (garden larger than 80sqm) (Ramsey et al., 2005); the presence of wild animals such as squirrel, opossum (Ramsey et al., 2005); and proximity to street lights (Dumonteil et al., 2013; Pacheco-Tucuch et al., 2012). The second group of risks pertains to the practice of keeping domesticated animals for subsistence: such as chickens in coops, rabbits, pigs; and for security: such as the presence of more than two dogs, (Dumonteil et al., 2013; Enger et al., 2004), and dogs able to enter and exit houses (Ramsey et al., 2005). The third group pertains to the levels of maintenance of peri-domestic areas such as cleaning trash from the peri-domestic area; the presence of rock piles (Dumonteil et al., 2013), presence of junk piles, and agricultural products (Enger et al., 2004).

Some of the risk factors were identified among two predominant types of peri-urban developments in South Texas. In colonias, environmental risks are a function of their small overall size which means that many residences are found in a peripheral location. A combination of regulatory and market conditions in colonias results in high rates of lot vacancies and housing abandonment, which means that many lots are surrounded by unconstructed or abandoned lots. Furthermore, because colonias lack municipal maintenance services or maintenance ordinances, many vacant lots may be overgrown with piles of trash and other material. On the other hand, lots in colonias are typically very small, a characteristic that is anticipated to decrease the risk of Chagas disease transmission.

Model Subdivisions are newer and typically larger peri-urban settlements than colonias, hence less dwellings are anticipated to be exposed on the periphery. The development of MSs has to comply with Model Subdivision rules which require the transfer of ownership of public streets, squares, and easements (among others) to a municipality. This means that right-of-way areas are maintained by a municipality, reducing the opportunity for wildlife to establish. On the other hand, MSs are platted into larger lots, which was identified as a risk characteristic. Furthermore, because Model Subdivisions are newer, some may have higher rates of unconstructed lots, also increasing the risk of Chagas disease. However, because the MS developments are younger, unconstructed lots are anticipated to have lower levels of overgrowth and littering than in colonias.

Overall, peri-urban developments in South Texas show many attributes that are anticipated risk factors for the establishment of Kissing bugs. The following section discusses the data and the methods that were used in this study to identify which of these risks explain the prevalence of Chagas disease in canines in the RGV region.

3 METHODOLOGY

This section starts with the research question and study objectives. Primary data that was provided for this research is described next, including an overview of how the data were acquired and resulting limitations for spatial analyses. Next, the geocoding process and the resulting spatial distribution of the sample are outlined. Then the section outlines the use of geographic information system (GIS) tools for secondary data collection from around each sampled canine. The last part details the steps of statistical analyses to explain the probability of canines testing positively for *T. cruzi* with environmental, land development, and socio-economic factors.

3.1 Research Question

This thesis attempts to answer the following research question:

What environmental, land development, maintenance, and socio-economic characteristics explain Chagas disease in canines in South Texas?

The answer to the research question will be determined through the analysis of the relationship between the prevalence of Chagas disease across locations in the RGV region and four groups of characteristics derived from the literature review: (1) Open Space characteristics: the types and sizes of surrounding cultivated (i.e. a pasture) and uncultivated land types (i.e. a forest); and distances to areas of protected wildlife (i.e. a state park); (2) Land Development characteristics such as the size and position within settlement, area covered by unconstructed lots, lot sizes, and lot uses; (3) Maintenance characteristics such as municipal maintenance; and (4) and socio-economic characteristics.

3.2 Prevalence of Chagas Disease in Canines

The current study builds on the results of a prior pilot research study that determined the prevalence of Chagas disease across seven locations in the RGV (Figure 4). Primary data (blood samples from canines) were collected by a team of researchers from Texas A&M University led by a Principal Investigator Dr. Sarah Hamer in 2015. The study was titled: *A multidisciplinary epidemiological approach to mitigate the human and animal health burden of Chagas disease across a transnational gradient*. I was involved with other stages of the study, including fieldwork and primary data collection in the summer of 2016, 2017, and a follow-up visit in the summer of 2018.

3.2.1 Institutional Review Board (IRB)

Both the current and the pilot studies were approved by the office of the IRB at Texas A&M University. The current study uses identifiable information (addresses) acquired during the pilot study. The Institutional Review Board number is: IRB2015-0279D and the expiration date is: 01/28/2020. I am a listed Research Assistant for the study.

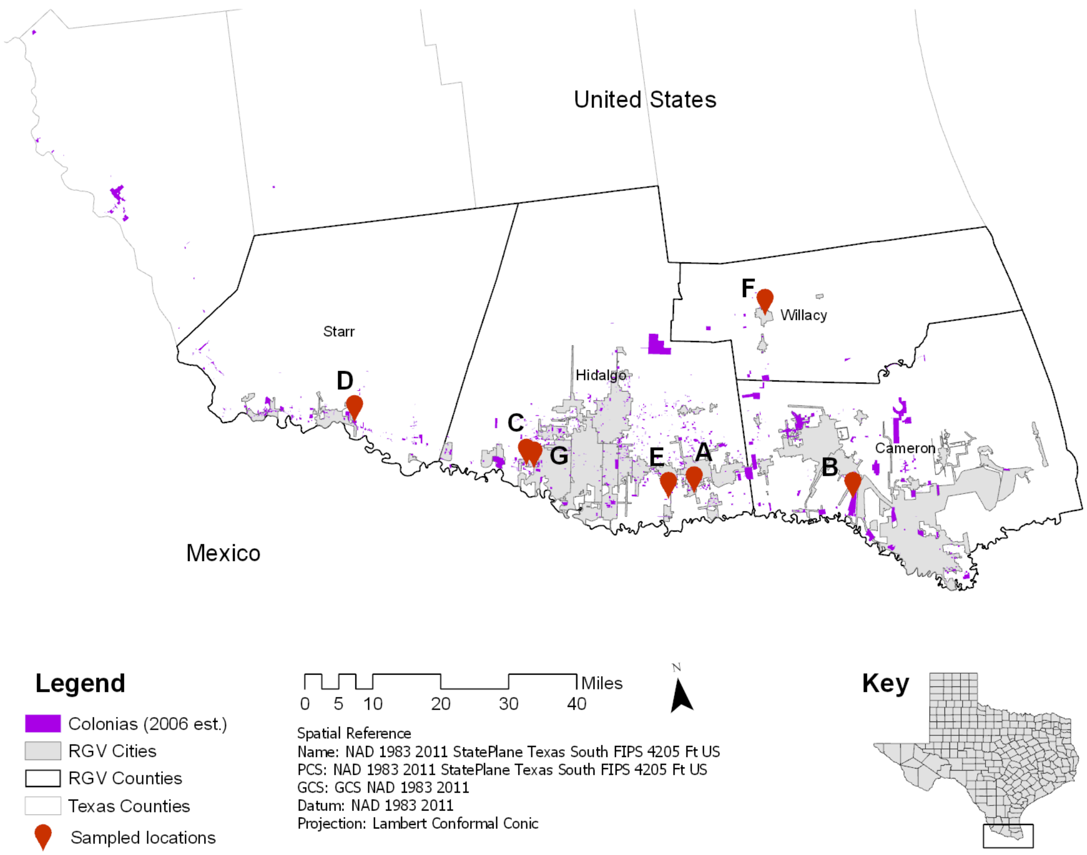

Figure 4: Location of sampled locations.
Data for RGV Cities, RGV Counties, and Texas Counties from U.S. Census Bureau (2016), and for Colonias from U.S, Geological Survey (Parcher & Humberson, 2007).

3.2.2 Pilot Study: Summary of Methods and Results

The pilot study was designed and executed to determine the prevalence of Chagas disease prevalence in humans and canines in the RGV region where higher risk of transmission was anticipated based on local conditions and previous studies. Blood samples were collected from humans and canines in summer 2015. The methods and results of both human and canine blood sample analyses are discussed in detail in an article by Curtis-Robles et al. (2017). Follows a summary of the study of canines only.

Blood samples were collected from the total of 209 canines and analyzed for the presence of *T. cruzi* antibodies, which were found in 41 canines (19.6 %). One of the seven locations that were sampled did not have any canines and was omitted from this current study (location G). The remaining six locations all had seropositive canines (Curtis-Robles et al., 2017). The team conducted a secondary test adjusting for false negatives which suggests even higher prevalence levels (31.6 %) in the region. A relationship between the location and seropositivity was confirmed by a bi-variate analysis (p= 0.130) using Fisher's exact test. A logistic regression revealed a significant (p=0.012) difference in the odds of seropositivity between locations F and A. Other factors pertaining to the breed, sex, or age of the canines did not show a significant difference. These results suggest that the geographic location and local conditions, including environmental and socio-economic ones, may play a role in the risk of Chagas disease transmission. Furthermore, the results present a considerable increase from previous studies of Chagas disease prevalence in South Texas and highlight the risk of transmission to humans in the region (Curtis-Robles et al., 2017). The primary data, including individual addresses of each sampled canine, and the results of the blood analysis were provided for this current study to assess local environmental conditions for association with seropositivity.

3.3 Geocoding

To assess characteristics surrounding each sampled canine, their addresses (provided by owners during initial blood sampling) were coded to a geographic database using an ArcGIS program. A database of all properties in the four counties was built using publicly available County Appraisal Districts (CAD)[3], and Texas Orthoimagery (2016) aerial photos. The CAD data set contains information about each property such as the address, name of the owner, lot size, lot use (residential, commercial, industrial, etc.), and other. The addresses provided by the owners and addresses in the CAD database were matched manually to minimize errors.

The following measures were taken to approximate incomplete or erroneous addresses: where house number could not be identified, the house number was

[3] Tax Appraisal data are publicly available on the website of each county, or at the County Clerk's office.

interpolated from adjacent known addresses; nearest house number was assigned; or the address of the canine owner was matched with the home owner listed in the CAD database. Where addresses could not be approximated, samples were omitted from the spatial analysis[4]. This process resulted in further reduction of the sample size, because some canines belonged to the same owner and thus the same address. The results of the geocoding process are described in the following section.

3.3.1 Geocoding Results

The 209 canines were assigned to 104 unique addresses, 4 of which could not be geocoded at all (3.8%), and 14 of which were approximated (13.5%). The resulting data set used for further analysis was reduced to n = 100 lots with at least one canine, and some lots with multiple canines (Table 1). 44 lots had only one canine per lot, 31 lots had two, 16 had three, and nine lots had between 4 – 7 canines.

---

[4] Some characteristics such as: location within the ETJ could be gleaned from the address, even though exact location could not be determined.

Table 1: Summary of the geocoding process.

Location	*Step 1: Sampled canines*			*Step 2: Unique addresses*			*Step 3: Geocoded lots*		
	Sampled	*Seropositive*		*Sampled*	*Seropositive canine or more*		*Sampled*	*Seropositive canine or more*	
		Count	*%*		*Count*	*%*		*Count*	*%*
A	22	7	31.8	8	4	50.0	8	4	50.0
B	54	12	22.2	30	9	30.0	30	9	30.0
C	14	4	28.6	5	2	40.0	5	2	40.0
D	32	7	21.9	18	7	38.9	17	7	41.2
E	23	5	21.7	10	4	40.0	10	4	40.0
F	64	6	9.4	33	6	18.2	30	6	20.0
G	0	0	N/A	0	0	N/A	0	0	N/A
TOTAL	209	41	19.6	104	32	30.8	100	32	32.0

Table 1 shows the three steps of the geocoding process in three columns. The Step 1 column shows the number of canines that were sampled in each of the six locations with code names A through to F (location G had no canines and was omitted from this current study). The range of canines per location was 14 – 64 and the seroprevalence per location ranged from 9.4% in location F to 31.8% in location A. The Step 2 column shows how many unique addresses were sampled within each location. The reduction in the number of unique addresses is due to the presence of multiple canines per lot, which was rather common across all locations. The seroprevalence was recalculated to represent at least one seropositive canine per unique address. The overall seropositivity per unique address ranged from 18.2% to 50%. Lastly, Step 3 shows the count and percentage of canines per lots that were successfully geocoded in ArcGIS. The final sample size in this study was n = 100 with an overall 32% of lots with at least one seropositive canine, and a range from 20% to 50% seroprevalence per location, and with a range of 5 to 30 lots per location. The following counts of lots were approximated or omitted in the process. In location B: 7 of 30 lots were approximated; D: 1 of 18 lots approximated, and 1 lot was omitted; and in location F: 6 of 33 lots were approximated, and 3 lots were omitted).

The geocoding process revealed the range of spatial distribution and clustering of sampled lots per county and per location. Figure 5 shows four of the six locations to illustrate the range of spatial clustering.

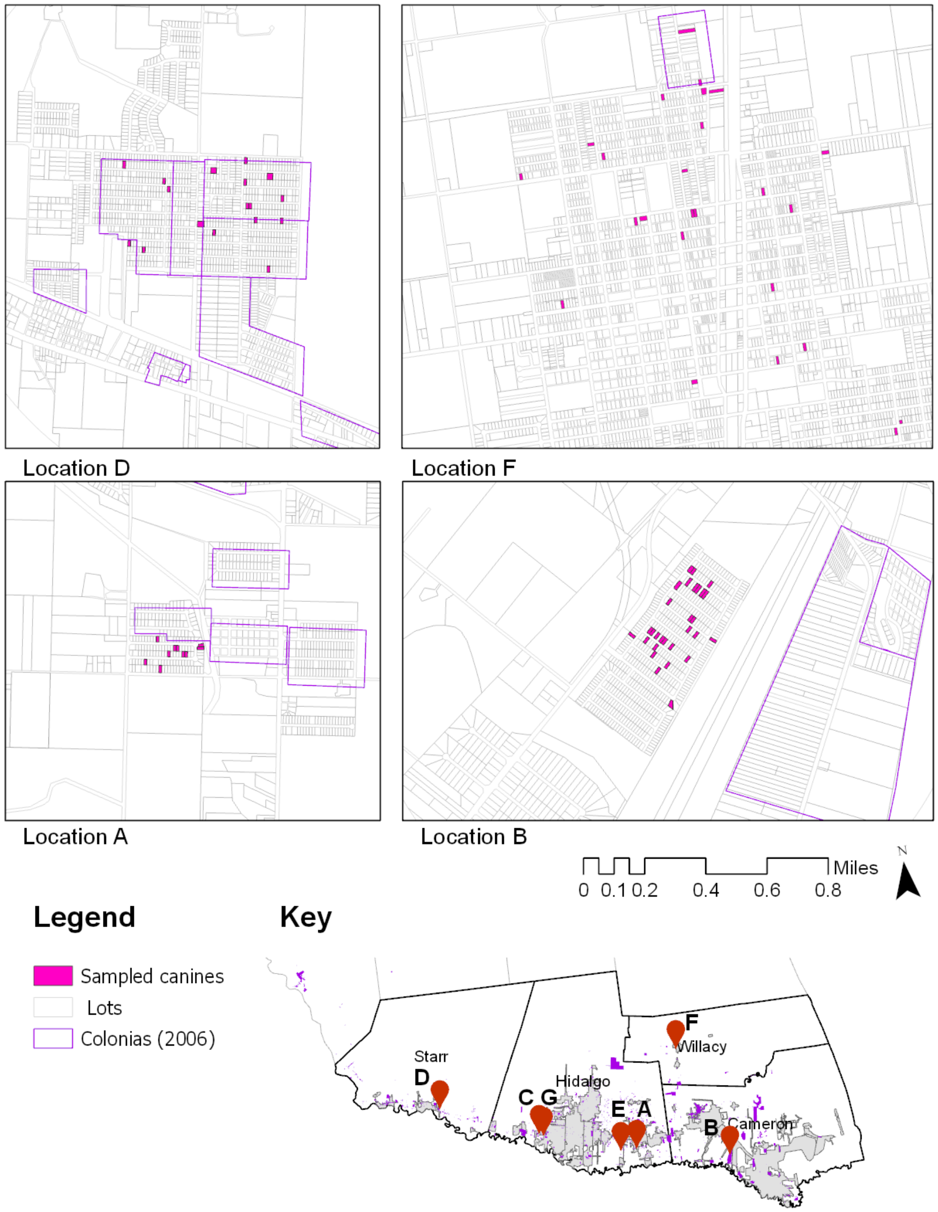

Figure 5: Geographic distribution of sampled lots across RGV counties.
Data for Lots from County Appraisal District for each county (Cameron Appraisal District, 2016 ; Hidalgo Appraisal District, 2017; Starr Appraisal District, 2017; Willacy Appraisal District, 2017), Data for RGV Cities, RGV Counties, and Texas Counties from U.S. Census Bureau (2016), and for Colonias from U.S, Geological Survey (Parcher & Humberson, 2007).

3.3.2 Geocoded Data Description

The sampled lots were spread across all four RGV counties: 30% in Willacy, 23% in Hidalgo, 30% in Cameron, and 17% located in Starr county. Nearly half of the lots were located within a Model Subdivision (49%), followed by a location in a city (28%), a colonia (19%), and 4% were located within an ETJ. The sampled lots were spread across six locations: A-F and generally clustered spatially within a neighborhood, although a variability of dispersal was detected. Locations A and B were both determined a Model Subdivision, and both were comparably smaller in size with respect to the area of developed land, with sampled canines clustered in several streets. Locations D and F were larger settlements with a more dispersed pattern of sampled lots across the space. Location D was a conglomeration of five colonias and canines were sampled in four of those colonias. Samples in location F belonged to a small town and an adjacent colonia.

3.3.3 Data Limitations

The data presented several methodological limitations for a spatial analysis of environmental conditions. First, even though the data collection was designed as a geographically stratified sampling to represent all four counties in the region, the spatial distribution within counties was concentrated in several geographic locations, and spatially clustered. Some of the locations were selected based on previous research contacts and reports of Chagas disease (Curtis-Robles et al., 2017). Snowball sampling was used in each location. As a result, the sample is not randomly distributed in space. This renders the sample both spatially biased and observations not independent, because those that are closer to each other have a higher chance of having similar socio-economic and environmental conditions. Secondly, the sample size is small (n=100), and prevents robust statistical analysis.

To address the spatial clustering and dependency of observations in the original study, spatial clusters were defined and included in the statistical analysis of the current study. The small sample size limits the external validity of the analysis, however an exploratory analysis may contribute towards the research design and data collection of future studies.

3.4 Secondary Data Source and Collection

A database of publicly available, secondary data on environmental, land development, socio-economic, and maintenance characteristics surrounding each sampled lot was compiled in ArcGISPro.

3.4.1 Secondary Data Sources

The following secondary data were collected from publicly available sources: County Appraisal District for each county of the RGV for the year 2016 or the year 2017 (Cameron Appraisal District, 2016 ; Hidalgo Appraisal District, 2017; Starr Appraisal District, 2017; Willacy Appraisal District, 2017); Aerial images (Texas Orthoimagery, 2016); U.S. Census municipal boundaries for 2016 (U.S. Census Bureau, 2016); American Community Survey 5 year estimate: 2013 -2017 Median household income in the past 12 months (in 2017 inflation-adjusted dollars) (2013-2017 American Community Survey 5-Year Estimates, 2017a); Texas State Park boundaries (*State Park Boundaries*, n.d.); Texas Wildlife Management Areas boundaries (*Wildlife Management areas* n.d.), Landcover data (U.S. Geological Survey, 2014); colonia outlines (Parcher & Humberson, 2007).

3.4.2 Secondary Data Collection

The bulk of the data were collected from within a 0.25mi buffer of each lot, because Kissing bugs are attracted by light from within that distance (Ryckman, 1981). In ArcGIS, a 0.25-mile circle was placed centered on the centroid of each lot to create a desired buffer. From within the buffer, spatial data were collected for the analysis. Land cover types were measured from high resolution Land cover raster data (cell size 30 x 30 m) (U.S. Geological Survey, 2014). The raster data were vectorized using a *raster-to-point* tool, creating spatial points (centroids) which are easier to select *by location* in ArcGIS than rectangular cells. Spatial points that intersected or were within the buffer outline were included. Total count for each buffer was 622 points, each representing an area of 900 square meter cell. These were summarized for each landcover type in each buffer using the *summarize* tool (Fig. 6).

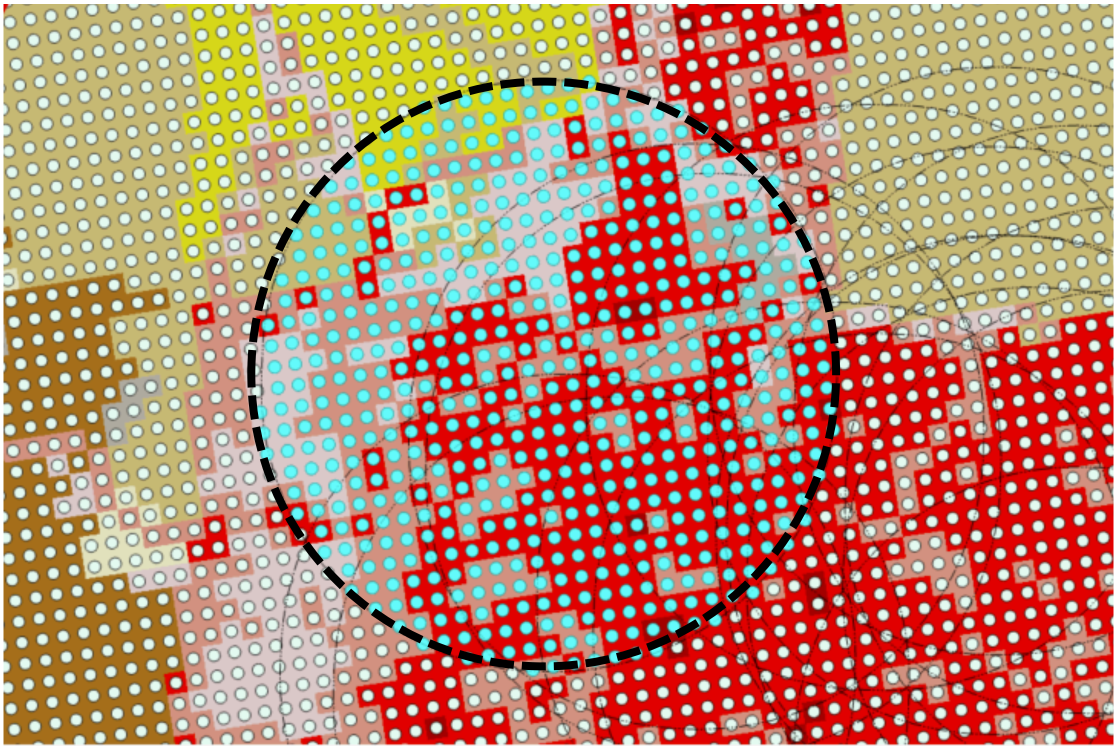

Figure 6: Land cover data with points overlaid for ease of selection for each buffer (i.e. dashed line circle) in ArcGIS program.
Data for Landcover Types from U.S. Geological Survey (2014).

Data pertaining to lot sizes were measured from Tax Appraisal data by selecting outlines of lots from within each buffer (Fig. 7). The selected outlines contained land development characteristics in associated tables. Lots that partially fell within the selection buffer were excluded if less than half of their area fell within the buffer, and vice versa.

Figure 7: Illustration of the selection of lots by size from within the quarter mile. Data for lots from Willacy Appraisal District (2017).

Similarly, data pertaining to lot use and vacancy were measured from Tax Appraisal data by selecting outlines of lots from within each buffer (Fig. 8). The selected outlines contained land development characteristics in associated tables. Residential lots (both occupied and vacant) within each quarter mile buffer were identified using CAD "state code" A, B, and C1, and C2 classifications,[5] and cross-checked with aerial photos.

[5] Texas Comptroller of Public Accounts (2014, p. 1) defines "A: Real Property: Single-family Residential; B: Real Property: Multifamily Residential; C1: Real Property: Vacant Lots and Land Tracts; C2: Real Property: Colonia Lots and Land Tracts"

CAD state code was only available for Cameron County, the remaining county data sets were updated manually from aerial photos to determine the land use as best as possible. This is a limitation of the study, a common limitation on the availability of secondary data in areas that may be considered "informal" by some institutions.

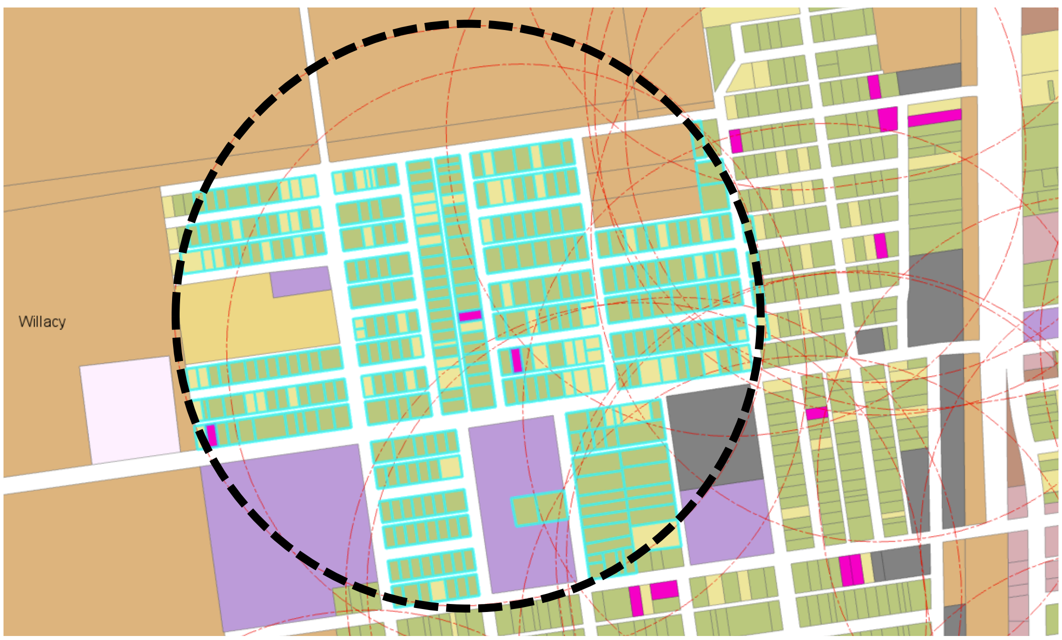

Figure 8: Illustration of the selection of lots by land use from within a quarter mile. Data for lots from (Willacy Appraisal District, 2017).

Socio-economic data were acquired in aggregate form from the Census Bureau at the block group level. Distance measurements to environmental features were measured manually (as the crow flies) in ArcGIS program to the nearest edge of a feature, in miles. Distances to state park was measured using Texas State Park boundaries (*State Park Boundaries*, n.d.), and distance to protected wildlife refuge was measure using Texas Wildlife Management Areas boundaries (*Wildlife Management areas* n.d.).

3.5 Operationalization of Variables

The selection of variables for this study was based on previous studies of environmental characteristic and Chagas disease. Because very few studies have been conducted in the United States, prior research studies in Mexico were reviewed to conduct exploratory analysis of potential variables. Even though climatic and environmental conditions in Mexico, particularly in northeastern states may be considered similar to Texas, as evidenced by niche modelling research studies. Land development practices differ significantly. Therefore variables were modified from available secondary data to suit the context of South Texas. Three broad groups of environmental characteristics are outlined below: Open space, Land development, and maintenance; one group of socio-economic characteristics, and one pertaining to canines.

The dependent variable is dichotomous and signifies whether a lot contains at least one canine that tested positively for *T. cruzi* (*positive lot*) or does not contain any (*negative lot*). Positive lots were coded “1” and negative lots were coded “0”.

3.5.1 Open Space Characteristics

Different types and size of cultivated and uncultivated open space (such as cropland, pasture, or forest) may play a role in the transmission of Chagas disease by offering habitat and harborage spaces for Kissing bugs and their hosts (Raghavan et al., 2015; Ramsey et al., 2005).

Raghavan et al. (2015) used National Land Cover Dataset set (NLDS) (*National Land Cover Database 2011* 2011) to examine this association. Building on the study, NLDS dataset was used to extract land cover types from within each buffer. Only open space land cover data were collected because more precise measurements of developed land were devised using tax appraisal data. The following 10 categories were present in the study area: Barren Land, Deciduous Forest, Evergreen Forest, Mixed Forest, Shrub/Scrub, Grassland/Herbaceous, Pasture/Hay, Cultivated Crops, Woody Wetlands, and Open Water. Due to low counts for some of these categories, several were combined into groups: Forest types were combined (Deciduous Forest, Evergreen Forest, Mixed

Forest) into the **ForestALL** variable. Uncultivated (**UncultivatedALL)** land cover types such as: Barren Land, Shrub/Scrub, Grassland/Herbaceous, Woody Wetlands, and Open Water were combined because they likely provide habitat opportunities for different types of wildlife than types of forests. Cultivated land cover types were retained ad separate variables (**Pasture** and **Crops)** to distinguish the impacts of pastures from crops. Pastures were expected to have a higher impact because domesticated animals, such as cattle kept overnight in enclosure, increase the source of bloodmeal for Kissing bugs. *The anticipated relationship is that a larger area covered by each type of open space will correspond with a higher prevalence of Chagas in sampled lots.*

Distance to a Natural Area (DistancePark)

Areas that provide protection for wildlife such as a state park or a wildlife refuge are considered reservoirs of Chagas disease. Additionally, waterbodies can serves as wildlife corridors. Where housing encroaches into proximity of wildlife habitat, a higher risk of peridomestic transmission is anticipated. Therefore, distances were measured to the nearest edge of a protected natural area (Texas State Park, Wildlife Refuge), or the riparian corridor of the Rio Grande and its tributaries, or a pond or a lake. *The* *anticipated relationship is that a shorter distance to a natural area will correspond with a higher prevalence of Chagas in sampled lots.*

3.5.2 Land Development Characteristics

Land development patterns pertaining to the size of settlement, position within the settlement (edge or the interior), vacancy rates, abandonment rates, lot sizes, and lot uses may support or hinder the establishment of an overlap between the sylvatic and the peri-domestic cycle of the disease by providing habitat. Three characteristics were identified to capture the extent and intensity of development surrounding each lot: Lot size, lot use, and vacancy rates.

In a study in Mexico, (Ramsey et al., 2005) identified that a garden size over 80 sqm (861 sqft) was associated with an increased risk of Chagas in Mexico, however lot sizes in South Texas are generally larger. 5,000 sqft is considered small in South Texas (Ward et al., 2004). To account for the lot sizes per buffer, areas consisting of large and small lots were included in the analysis.

Area of Small Lots

Small lot size variable was established to examine whether more densely platted development pattern is associated with lesser risk of Chagas disease. Because the sample is mainly located in colonias, and Model Subdivisions, small lot size was determined from the literature on colonias and Model Subdivisions. Ward et al. (2004) state that 5,000 sqft is a small lot in the South Texas region. Durst (2016) shows that median sizes of lots in colonias are smaller than in Model Subdivisions (approx. 8,700 and 21,800 sqft, respectively); he lists median sizes of lots in Cameron, Hidalgo, and Starr counties are between 6,100 and approx. 8,300 sqft., where else in MSs between approx. 21,800 and 25,300 sqft. Small size lot was defined at 7,000sqft for the purpose of the study. This definition includes standard sizes as follows: 50x115ft (5,750 sqft), 50x120ft (6,000 sqft), 60x100ft (6,000 sqft), 60x115ft (6,900 sqft), and 70x100ft (7,000 sqft), and variations under 7,000 sqft. To allow for discrepancy and to include lots of that size, the exact cut off area was set at 7,005 sqft in ArcGIS program. The data were derived from CAD shapefiles by selecting lots that were located within, or intersected the quarter mile buffer and subtracting lots larger than 7,005 sqft. *The anticipated relationship is that* *more area covered by small lots will correspond with lower prevalence of Chagas in sampled lots.*

Area of Large Lots

Area covered by lots larger than half an acre (21,780 sqft) within each buffer was included as a variable. The size corresponds with the minimum allowable lot size for septic tank sewerage disposal according to the ordinance of the city of Brownsville (Brownsville, 2018). This distinction captures a major difference in development patterns. The data were derived from CAD shapefiles by selecting all lots that intersected the quarter mile buffer, discounting lots larger than 21,780 sqft, and finally subtracting this figure from the area of the buffer. This inversion of the selection process was to eliminate steps such as clipping large lots to the boundary, and recalculating their area. Lots that were partially included in the buffer were excluded if less than half of their area fell within the buffer. *The anticipated relationship is that more area covered by lots larger than half acre will correspond with higher prevalence of Chagas in sampled lots.*

Area of Residential Lots

Residential use variable was established to assess the relationship between land uses and the risk of Chagas disease. Residential lots (both occupied and vacant) within each quarter mile buffer were identified using CAD "state code" A, B, and C1, and C2 classifications[6]. *The anticipated relationship is that more area covered by residential will correspond with lower prevalence of Chagas in sampled lots.*

Edge/ETJ

A peripheral location of lots within a settlement was identified as a risk characteristics in Mexico (Ramirez-Sierra et al., 2010), where the location of 200m or more from the edge of a settlement had significantly lower level of infestations by Kissing bugs than the edges of the settlement. To fit the lower density of settlement in South Texas in comparison with Mexico, 400m distance was established to demarcate an Edge zone by offsetting census shapefiles of places within the study area. Lots were

---

[6] Texas Comptroller of Public Accounts (2014, p. 1) defines "A: Real Property: Single-family Residential; B: Real Property: Multifamily Residential; C1: Real Property: Vacant Lots and Land Tracts; C2: Real Property: Colonia Lots and Land Tracts"

assigned into two categories: within the center of a settlement = “0” and within the Edge of the ETJ = “1”. *The anticipated relationship is that location within the Edge/ETJ will correspond with higher prevalence of Chagas in sampled lots.*

3.5.3 Maintenance Characteristics

The larger and less used the lot is, the less private maintenance of the lot such as, and the more opportunity for Kissing bugs and their hosts to establish a nest in the vicinity of domestic animals and humans. Additionally, municipal maintenance varies across subdivision types with respect to incorporation within city limits, colonia, or a Model Subdivision. The following explanatory variables were established to explain the risk of Chagas disease by level of private and municipal maintenance.

Lot Size

Ramsey et al. (2005) found that large lot (with a garden of 80 sqm and more) were a risk factor in Mexico because of the reduced ability to maintain the larger garden. Lot size of each sampled lot was established as a variable to account for the size of the individual lots as a risk characteristic. *The anticipated relationship is that larger area per sampled lot will correspond with higher prevalence of Chagas in sampled lots.*

Area of Uncompromised Lots

A variable combining size and vacancy rates was established. A lot was deemed uncompromised, if the lot was smaller than 7,005 sqft and not vacant. The sum of areas of uncompromised lots was calculated for each buffer by subtracting the area of Unconstructed lots from the **Area of Small lots**. Unconstructed lots within the buffer were identified using the steps described in the following variable description: Adjacent Unconstructed lots, except for the whole buffer area. *The anticipated relationship is that more area covered by uncompromised lots will correspond with lower prevalence of Chagas in sampled lots.*

Adjacent Unconstructed Lots

Ramsey et al. (2005) found that an increase in unconstructed lots in immediate adjacency, increased the odds of Chagas disease in Mexico. Vacancy rates and abandonment are a problem in colonias (Durst & Ward, 2015). Unconstructed lots were identified using CAD "state code" C1 classification,[7] and cross-checked with Texas

---

[7] Texas Comptroller of Public Accounts (2014, p. 1) defines "C1: Real Property: Vacant Lots and Land Tracts"

Orthoimagery (2016) aerial photos manually by observing eight adjacent lots (left and right sides, front and back, and four corner lots). *The anticipated relationship is that a higher count of unconstructed lots adjacent to sampled lots will correspond with higher prevalence of Chagas in sampled canines.*

Subdivision Types

The lack of municipal or private maintenance and resulting presence of junk piles and overgrown vegetation provides harborage spaces for kissing bugs and their hosts. Municipal maintenance varies across settlements depending on incorporation status: City, Colonia, MS, ETJ. *The anticipated relationship is that higher levels of maintenance (location within a city) will correspond with lower prevalence of Chagas in sampled canines.*

3.5.4 Socio-economic Status

Socio-economic status was identified as an overarching factor in the spread of Chagas disease (Hotez et al., 2012). In the absence of individual level data, the following characteristics were used as a proxy for socio-economic status in this study (Table 2).

Median Household Income

Block group Median Household Income was selected as the smallest measure of socio-economic status in the absence of Tax Appraisal house value data in three of the four counties. American Community Survey estimates for the year 2017 values were used (2013-2017 American Community Survey 5-Year Estimates, 2017b). Three block groups one settlement in Willacy county had missing values and were replaced with average values for that settlement. *The anticipated relationship is that lower Median HH income will correspond with higher prevalence of Chagas in sampled canines.*

County

A categorial variable representing location within one of the four counties was established as a variable to capture the differences in measures of poverty across the four counties. *The anticipated relationship is that higher levels of poverty will correspond with higher prevalence of Chagas in sampled canines.*

Table 2: Candidate variables and their expected impact

Group	*Candidate Variable*	*Explanation*	*Exp'd impact*
Dependent Variable			
	Positive/ Negative Lot [binary]	At least one canine in a lot tested positively for *T.cruzi* = 1, None = 0	
Independent variables			
Open Space	UncultivatedALL [cell]	Sum of cells within 0.25-mile	+
	Pasture [cell]	Sum of cells within 0.25-mile	+
	Crops [cell]	Sum of cells within 0.25-mile	+
	ForestALL [cell]	Sum of cells within 0.25-mile	+
	Distance to Park [mi]	Euclidean distance to nearest park / refuge	-
Land Develo pment	Large Lots [sqft]	Sum of lots ≥ 0.5acre within 0.25-mi	+
	Small Lots [sqft]	Sum of lots ≤ 7,000sqft within 0.25-mi	-
	Residential Lots [sqft]	Sum of residential/vacant lots within 0.25mi	-
	Edge/ETJ [binary]	Lot in ETJ or edge of settlement ='1'; lot within center of settlement = '0'	+
Mainte nance	Subdivision Types [categ]	Location within a type of subdivision	
	Lot Size [sqft]	Size of sampled lot	+
	Adjacent Unconstructed [count]	Count of unconstructed lots adjacent to each	+
	Uncompromised Lots [sqft]	Sum of Uncomprised lots (≤ 7,000 sqft & occupied) within 0.25-mile buffer	-
Socio-econ	Median HH Income [$]	Location within block group	-
	County [category]	Location within RGV county	
Canines	No. of Canines [count]	Count of canines per lot	+

3.6 Statistical Analysis

Statistical analyses used in this study were a bivariate and a multivariate binomial logistical regression analyses. The former was used to screen candidate independent variables for their explanatory power and to decide which variables to include in the multivariate analysis. Fisher's exact test was used. The multivariate logistical regression was used to explain the variance in the dichotomous dependent variable based on multiple explanatory variables, which can be continuous or nominal. STATA program (Stata/IC 15.1 for Mac 64-bit Intel, Revision 08 Mar 2018) was used for these analyses.

3.6.1 The Logistical Regression Model

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots \beta_m x_m$$

Where: π indicates the probability of a lot with at least one canine with *T.cruzi*,

β_0 is the intercept coefficient

$x_1 .. x_m$ are the explanatory variables

$b_1, b_2, b_3, .. b_k$ are the coefficients for the explanatory variables

3.6.2 Assumptions of Logistical Regression

Several assumptions apply to a logistical regression analysis (Kassambara, 2018; UCLA Institute for Digital Research and Education, n.d.), which must be met for the logistical regression to provide valid results. The following section describes the assumptions and how they were addressed in this study.

Dichotomous Dependent Variable: Binomial logistic regression requires that the dependent variable be dichotomous: has two categorical, independent groups, and the two categories of the mutually exclusive. In this study, the dependent variable is dichotomous, because the sampled lot either contains at least one canine that tested positively for *T. cruzi* (*positive lot*) or does not contain any (*negative lot*). This assumption was met and logistic regression was selected as the appropriate statistical test for this study.

Independent Observations: Logistic regression requires that observations are independent of each other (data should not be repeated measurements or matched data). Based on the character of the sampling of original data, the dataset is spatially clustered, therefore observations were not independent. To compensate for the clustering, nine clusters were explicitly defined based on geographic location. Clusters were identified using the quarter-mile buffer to determine whether or not lots were in proximity to each other so that Kissing Bugs could be attracted from one lot to the next. Where quarter mile buffers of sampled locations overlapped, those lots were considered part of the same spatial cluster (Fig. 9). In other cases, lots were located at a distance from other lots such as in the case of location B, where one sampled lot was located several miles south west from the rest of the sampled lots. (Fig. 9). Following the quarter mile cut off distance, the data set was categorized into nine spatial clusters, with locations A, B, and F split into two clusters each, and the remaining locations C, D, and E representing a spatial cluster each. Cluster variable was included in the statistical analysis using the *cluster* function.

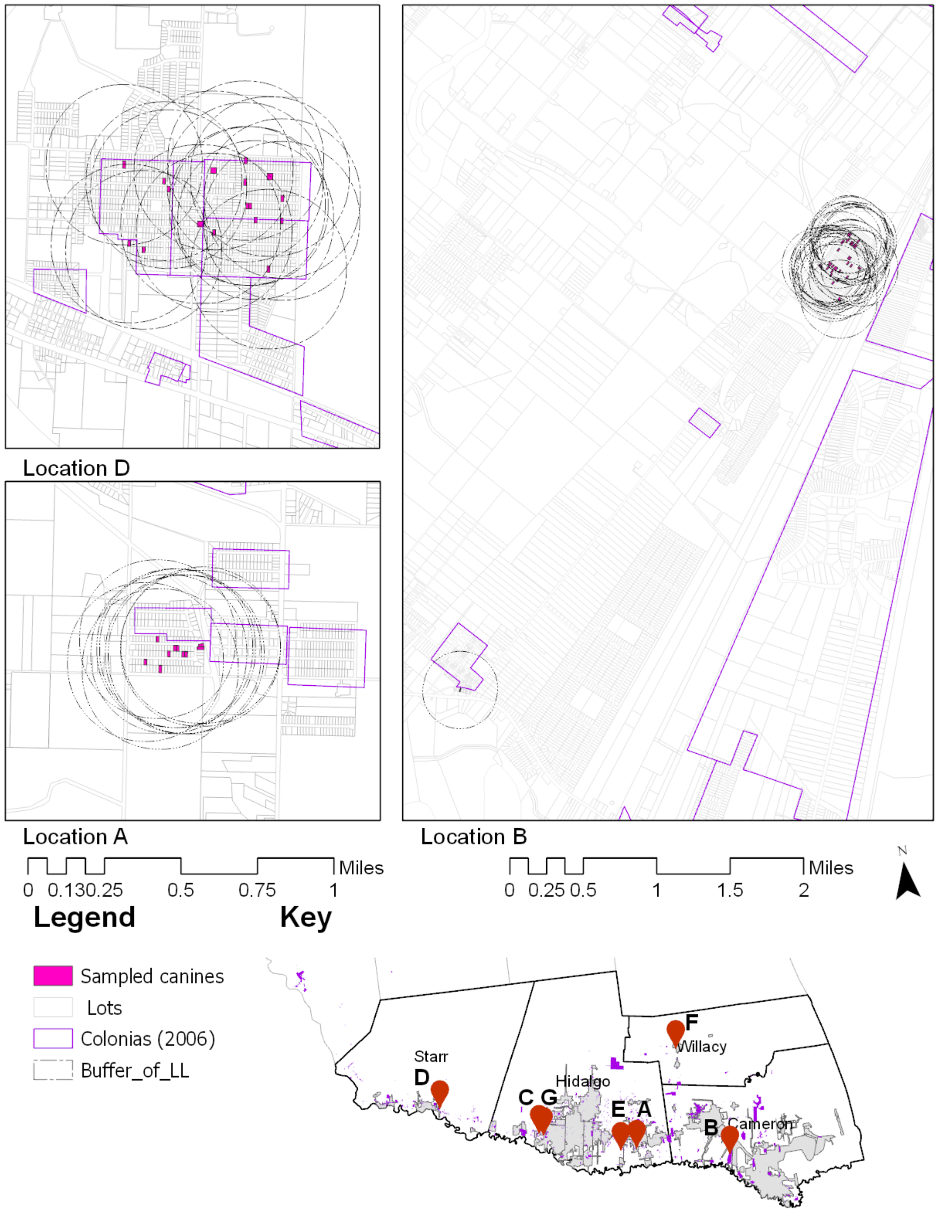

Figure 9: Examples of quarter mile buffers overlaps and locations.
Data for Lots from County Appraisal District for each county (Cameron Appraisal District, 2016 ; Hidalgo Appraisal District, 2017; Starr Appraisal District, 2017; Willacy Appraisal District, 2017), Data for RGV Cities, RGV Counties, and Texas Counties from U.S. Census Bureau (2016), and for Colonias from U.S, Geological Survey (Parcher & Humberson, 2007).

Low Multicollinearity: Logistical regression is sensitive to multicollinearity among independent variables. Variables that are highly correlated with each will result in large (inflated) standard errors for coefficients and the estimated coefficients for the regression can be unreliable (UCLA Institure for Digital Research and Education, n.d.). To assess collinearity among independent variable, tolerance measure was calculated for each variable using the Collinearity Diagnostic Tool (Collin)[8] in STATA. Tolerance levels higher than 0.2 are concerning and should be addressed (Menard, 2010). This assumption was satisfied by calculating tolerance measures for all independent variables (Table 3). Several variables were excluded because the levels of tolerance were below 0.2 level: **SmallLots**, **TypeSubdiv**, **County**. The latter two variables were collinear due to sampling and will be retained for bivariate analysis. Within the multivariate analysis, these variables will be replaced with the cluster function to account for clustering. One additional variable was close to 0.2 tolerance level: **LargeLots**, and should be interpreted with caution. (Condition No.: 44.3364).

[8] Collin tool can be downloaded from: https://stats.idre.ucla.edu/stata/ado/analysis/

Table 3: Result of collinearity analysis.

Variable	*VIF*	*SQRT VIF*	*Tolerance*	*R-Squared*
LotPos	1.58	1.26	0.6341	0.3659
ForestALL	1.61	1.27	0.6212	0.3788
UncultivatedALL	2.17	1.47	0.4609	0.5391
Pasture	1.68	1.3	0.5959	0.4041
Crops	2.4	1.55	0.4167	0.5833
DistForest	2.5	1.58	0.3993	0.6007
LotHalfAcre	4.51	2.12	0.2217	0.7783
ResiOnly	2.85	1.69	0.3511	0.6489
EdgeETJ	2.14	1.46	0.4668	0.5332
PercUncomprLots	3.59	1.89	0.2786	0.7214
AdjUnconstr	1.44	1.2	0.695	0.305
MedianHHincome	1.97	1.4	0.5084	0.4916
NoCanine	1.34	1.16	0.7461	0.2539
Mean VIF	2.29			

Linearity: Logistic regression requires that there is a linear relationship between each continuous independent variable and the logit of the dependent variable. I used the Box-Tidwell test in STATA to assess linearity of independent variables. The "boxtid"[9] function was downloaded and installed for STATA. This assumption was violated by all continuous variables. To minimize the violation, continuous variables were recoded into categorical variables.

---

[9] Boxtid downloaded from: https://stats.idre.ucla.edu/stata/ado/world/

ForestALL, **UncultivatedALL**, **Pasture**, and **Crops** were recoded to categorical variables of the percentage of coverage of each landcover type within the quarter mile buffer. All four variables were zero-inflated: many cells contained a zero because landcover types were not present in many sampled locations. For this reason, a category of 0% = 0 was established for all. The rest of the categories were established with respect to the distribution of each variable, based in quartile increments. **ForestAll** was had a minimum of 0% and maximum of 20.6% coverage across the sample. Two categories were established: 0% = 0; 0.01 – 25% = 1. **UncultivatedALL** variable had a minimum of 0% and a maximum of 44.7%. The variable was reclassified into three categories: 0% = 0; 0.01 – 25% = 1; 25.01 – max = 2. **Pasture** variable had a minimum of 0% and a maximum of 55.3%. The variable was reclassified into four categories: 0% = 0; 0.01 – 25% = 1; 25.01 – 50% = 2; 50.01 – max = 3. **Crops** variable had a minimum of 0% and a maximum of 60.45%. The variable was reclassified into four categories: 0% = 0; 0.01 – 25% = 1; 25.01 – 50% = 2; 50.01 – max = 3.

DistPark was recoded to a dichotomous variable with two categories: Less than one mile (5,280 ft) distance and more than one mile distance. One mile was selected to approximate an area of risk, because the maximum flight range of a Kissing bug is one mile (Wood & Anderson, 1965). 0 – 5,280ft = 1; and 5,280.01 and more = 0.

LargeLots, ResiLots, UncomprLots were all recoded to categorical variables with quartiles as categories: 0 - 25% = 1; 25.01 – 50% = 2; 50.01 – 75% = 3; 75.01 – 100% = 4.

MedianHHincome was recoded to a categorical variable with three categories based on the following thresholds: poverty line for four person household, median household income for RGV counties, and median household income for Texas. Poverty line for a four person household was $25,094 in 2017, and applies to the whole of U.S. (U.S. Census Bureau, 2017). Four person household was selected because the average household size in the Rio Grande Valley counties ranged from 3.4 – 3.8 in 2017 (U.S. Census Bureau, 2017). Median household income for RGV counties was calculated as the median of 2017 values for each county in the RGV: (Cameron: $36,095, Hidalgo: $37,097, Starr: $27,133, Willacy: $29,104) = $32,600 (2013-2017 American Community Survey 5-Year Estimates, 2017b). Median household income for Texas estimate for 2017 was 57,051. None of the block groups in the RGV surpass the value of the Texas median household income. The MedianHHincome variable was hence recoded into three categories: less than 25,094 = 1; 25,094.01 – 32,600 = 2; and 32,600.01 - 57,051 = 3, all in 2017 dollars adjusted for inflation (2013-2017 American Community Survey 5-Year Estimates, 2017b).

Table 4: Definition of variables.

Group	Variable	Definition
Open Space	ForestALL	Binary [%] 0 = 0; 0.01 – 20.6 (max) = 1;
	UncultivALL	Categorical [%] 0 = 0; 0.01 - 25 = 1; 25.01 – max = 2
	Pasture	Categorical [%] 0 = 0; 0.01 - 25 = 1; 25.01 - 50 = 2; 50.01 – max = 3
	Crops	Categorical [%] 0 = 0; 0.01 - 25 = 1; 25.01 - 50 = 2; 50.01 – max = 3
	DistPark	Binary [ft] 0 – 5,280 = 1; 5,280.01 - max = 0
Land Dvlp.	LargeLots	Categorical [%] 0 - 25 = 1; 25.01 - 50 = 2; 50.01 - 75 = 3; 75.01 - 100 = 4
	ResiLots	Categorical [%] 0 - 25 = 1; 25.01 - 50 = 2; 50.01 - 75 = 3
	Edge/ETJ	Binary Lot in ETJ or within the edge of settlement = 1; lot within the center of settlement = 0
Maintenance	SubdType	Categorical City = 1; Colonia = 2; ETJ = 3; MS = 4
	AdjUnconstr	Categorical [lots] 0 = 0; 0.5 - 1 = 1; 1.5 - 2 = 2; 3 = 3; 4 = 4
	UncompLots	Categorical [lots] 0 - 25 = 1; 25.01 - 50 = 2
Socio-econ	MedianHH Income	Categorical [$] min – 25,094 = 1; 25,094.01 – 32,600 = 2; 32,600.01 – 57,051 = 3
	County	Categorical [county] Willacy = 1; Starr = 2; Cameron = 3; Hidalgo = 4
Canines	NoCanines	Count [canines] 1 = 1; 2 = 2; 3 = 3; 4/max = 4

4 RESULTS

4.1 Descriptive Results

Descriptive characteristics of the sample by open space, land development, socio-economic, maintenance, and canine characteristics are presented in the following sections, each detailing a group of variables.

Open Space group: Majority of sampled lots had only up to a quarter of the land of their buffer covered by any of the Open Space landcover type (Forests; Uncultivated; Pasture; and Crops). In other words, much of the area of each buffer was covered by developed land. The least represented landcover type was the forest. With regards to the distance to a natural area (i.e. distance to a state park, wildlife refuge, or the riparian corridors of rivers and water bodies), approximately two thirds of all cases were located further than 5,280 ft, and approx. one third of cases were located in proximity to at least one of these natural areas.

Land Development group: The majority (63%) of sampled lots had up to half of their quarter mile buffers covered with large lots. 15 cases had more than half of each lots covered by large lots. Similarly, the majority of cases (77%) had between a quarter buffer to a half a buffer covered by residential lots. In 16 cases, more than half of the cases' buffers were covered by residential lots. Of all cases, majority (87%) were located on the edge of a settlement or in the extraterritorial jurisdiction, with the remaining 13% located in the central zone of a settlement.

Maintenance group: Almost half of sampled lots were located within a Model Subdivision (49%), followed by a location in a city, or a colonia (28 and 19 cases, respectively), and only 4 cases were homesteads in the extraterritorial jurisdiction. Over half (64%) of the sampled lots had at least one unconstructed lot in their adjacency. Most cases (69%) had up to a quarter of their buffer comprised of uncompromised lots, with approximately a third of the cases (31%) having a larger proportion of their buffer covered by uncompromised lots.

Socio-economic group: Approximately third of the sampled lots were located in block groups with median household income below the poverty line. Half of the lots were located in block groups with median household income above the poverty line but below the RGV median, and the remaining 20% had a higher median income than the median in RGV, but lower than the median in the State of Texas. The sample was spread across all four counties in the RGV. Willacy and Cameron counties both had a third of the sampled lots, with Starr and Hidalgo conties both containing approximately 20% of the sampled lots.

Canine group: Nearly half of all lots (45%) had one canine present, a third (30%) of lots had two canines, and a quarter (25%) of lots had more than three canines present.

4.2 Results of Bivariate Analysis

Fisher's exact test was used to assess the statistical significance of the difference in characteristics among positive and negative lots. Results with a *p-value* ≤ 0.25 were selected for inclusion in the logistical regression analysis. One each variable pertaining to open space, land development, maintenance, and canine characteristics showed evidence of significant difference among positive and negative cases (Tables 5 - 9). None of the socio-economic, and landcover type open space variables showed a statistically significant difference.

Among the Open Space group of variables (Table 5), only the Distance to Park (a state park, wildlife refuge, or the Rio Grande riparian corridor) variable showed a significant result. Among positive cases approximately double the proportion of lots were located closer than 5,280 ft from the natural area (48%) as opposed to being located further (25%). No other variable in this group showed a significant result.

Table 5: Results of bivariate analysis for Open Space variables.

Variable	*Total sample N = 100*	*Negative lots*	*Positive lots*	*p-value*
Open Space:				
ForestALL (%)				
0	52	37 (53.6)	15 (48.4)	0.670
0.01 – 25	48	32 (46.4)	16 (51.6)	
Uncultivated (%)				0.937
0	12	9 (13.0)	3 (9.7)	
0.01 – 25	75	51 (73.9)	24 (77.4)	
25.01 – 44.7 (max)	13	9 (23.0)	4 (12.9)	
Pasture (%)				0.489
0	42	28 (40.6)	14 (45.2)	
0.01 – 25	43	30 (443.5)	13 (41.9)	
25.01 – 50	14	11 (15.9)	3 (9.7)	
50.01 – 55.3 (max)	1	0 (0.0)	1 (3.2)	
Crops (%)				0.499
0	36	25 (36.2)	11 (35.5)	
0.01 – 25	22	17 (24.6)	5 (16.1)	
25.01 – 50	27	19 (27.5)	8 (25.8)	
50.01 – 60.5 (max)	15	8 (11.6)	7 (22.6)	
Distance to Park (ft)				0.030
More than 5,280	73	55 (79.7)	18 (58.1)	
Less than 5,280	27	14 (20.3)	13 (41.9)	

Among the Land Development group of variables, none of the variables showed a significant result below the cut off level in a two-sided test (Table 6). However the Edge/ETJ variable showed a result below the cut off level in a 1-sided test. Because the anticipated relationship was that cases located within the settlement center will be at a lower risk of Chagas disease transmission, and because of the support from literature review, this variable was included in the multiple regression analysis. Among positive cases, 6% were located within the settlement center, in comparison with the majority of positive cases (94%) located within the ETJ or within the edge of a settlement.

Table 6: Results of bivariate analysis for Land Development variables.

Variable	*Total sample N = 100*	*Negative lots n (%)*	*Positive lots n (%)*	*p-value*
Land Development:				
Large Lots (%)				0.522
0 – 25	22	15 (21.7)	7 (22.6)	
25.01 – 50	63	45 (65.2)	18 (58.1)	
50.01 – 75	14	9 (13.0)	5 (16.1)	
75.01 – 100	1	0 (0.0)	1 (3.2)	
Residential Lots (%)				0.796
0 – 25	7	4 (5.8)	3 (9.7)	
25.01 – 50	77	54 (78.3)	23 (74.2)	
50.01 – 75	16	11 (15.9)	5 (16.1)	
Edge/ETJ location			1-sided test:	0.163
Within settlement center	13	11 (15.9)	2 (6.45)	
On the edge or in ETJ	87	58 (84.1)	29 (93.6)	

The bivariate analysis for the Maintenance variables showed significant results below the cut off level for the Adjacent Unconstructed variable (Table 7). Among positive cases, the majority (71%) had between one half and two adjacent lots unconstructed, where else almost half of negative cases had no unconstructed lot in their adjacency. This result was significant at the 0.05 *p-value* level. Among Subdivision types, the highest proportion of positive cases was in Model Subdivisions (55%), followed by Colonias (23%), and the city (21%). This relationship did not show statistical significance.

Table 7: Results of bivariate analysis for Maintenance variables.

Variable	*Total sample N = 100*	*Negative lots n (%)*	*Positive lots n (%)*	*p-value*
Maintenance:				
SubdivType				0.575
City	28	22 (31.9)	6 (19.4)	
Colonia	19	12 (17.4)	7 (22.6)	
ETJ	4	3 (4.4)	1 (3.2)	
MS	49	32 (46.4)	17 (54.8)	
AdjUnconstr				0.039
0	36	30 (43.5)	6 (19.4)	
0.5-1	41	26 (37.7)	15 (48.4)	
1.5-2	19	12 (17.4)	7 (22.6)	
3	2	0 (0.0)	2 (6.5)	
4	2	1 (1.5)	1 (3.2)	
Uncompromised Lots (%)				1.000
0 – 25	69	48 (69.6)	21 (67.7)	
25.01 – 50	31	21 (30.4)	10 (32.3)	

In the socio-economic group of variables (Table 8), no results showed a statistical significance below the *p-value* cut off level, although County variable was approaching the cut off level. The Median HH income between the poverty line ($25,094) and the median household income in the region ($32,600) had the highest rate of positive cases (58%), followed by income below the poverty line with 8 positive cases (26%), and 5 positive cases in the higher income bracket (16%). The sample was spread across all four counties. Among positive lots, the most were located in Hidalgo and Cameron counties (32% and 29% respectively). The County variable was not included in the multivariate analysis as it was collinear with many other predictor variables.

Table 8: Results of bivariate analysis for Socio-economic variables.

Variable	*Total sample N = 100*	*Negative lots n (%)*	*Positive lots n (%)*	*p-value*
Socio-economic:				
MedianHH Income ($)				0.627
Less than 25,094	31	23 (33.3)	8 (25.8)	
25,094 – 32,600	50	32 (46.4)	18 (58.1)	
32,600 – 57,051	19	14 (20.3)	5 (16.1)	
County				0.317
Willacy	30	24 (34.8)	6 (19.4)	
Starr	17	11 (15.9)	6 (19.4)	
Cameron	30	21 (30.4)	9 (29.0)	
Hidalgo	23	13 (18.8)	10 (32.3)	

The results for the number of canines per lot showed evidence of highest significance among all predictor variables at the *p-value* = 0.000 level. The highest rate among positive cases was in lots with two canines (32%), followed by lots with four or more canines (26%). Among negative cases, the highest rate (55%) was for single canine cases, followed by lots with canines (29%), three canine lots (15%), and only one case with four or more canines.

Table 9: Results of bivariate analysis for the number of canines per lot.

Variable	*Total sample N = 100*	*Negative lots n (%)*	*Positive lots n (%)*	*p-value*
Canines:				
No of canines per lot				0.000
1	45	38 (55.1)	7 (22.6)	
2	30	20 (29.0)	10 (32.3)	
3	16	10 (14.5)	6 (19.4)	
4 or more	9	1 (1.5)	8 (25.8)	

4.3 Results of Multiple Logistical Regression Analysis

Results of the multiple logistical regression are shown in Table 10 with odds ratios, standard errors, z – scores, significance levels (*p-values*) and 95% confidence intervals for each independent variable. Values of more than one for the odds ratio per given independent variable mean an increase in the odds of positive lots (at least one canine tested positively for *T.cruzi*) and values of less than one mean a decrease in the odds of positive lots for a given independent variable. From the bivariate screening, four explanatory variables were included in the building of the final model, two of which were included as dummy variables. To compare models, several measures of fit were computed using the "statfit" command. A reduction in the AIC and the BIC measures was considered an improvement to the model (Williams, 2018). Three explanatory variables were included in the final regression based on the model fit measures. Edge/ETJ variable, and the cluster function were omitted from the final model. The Edge/ETJ variable increased the measures of fit, and the clusters were not of a sufficient rank to perform the model test. This was likely due to the low cell counts for some of the isolated, small clusters.

Table 10: Results of logistic regression model for *T. cruzi* prevalence in lots.

Dependent variable: LotPos	*Odds Ratio*	*Std. Err*	*z*	*p-value*	*95% Conf. Int.*	
Distance to Park	2.697	1.575	1.7	0.089	0.859	8.470
Adjacencent Unconstr (0)				Referent		
Adjacencent Unconstr (0.5-1)	4.826	3.260	2.33	0.020	1.284	18.139
Adjacencent Unconstr (1.5-2)	2.718	2.122	1.28	0.200	0.588	12.558
Adjacencent Unconstr (4)	11.752	18.983	1.53	0.127	0.496	278.633
Number of canines (1)				Referent		
Number of canines (2)	3.178	2.083	1.76	0.078	0.880	11.481
Number of canines (3)	3.207	2.259	1.65	0.098	0.806	12.759
Number of canines (4+)	76.670	96.033	3.46	0.001	6.584	892.878
Constant	0.048	0.034	-4.3	0.000	0.012	0.190

The final model explained approximately 23% of the variance in the dependent variable (Pseudo r-square = 0.2336). Several variables were significant at the 0.1, 0.5, and 0.1 *p-value* levels. The odds of a positive lot (*T. cruzi* in lot) were nearly three times higher (270%) in lots located within 1-mile of a natural area (wildlife refuge, state park, etc.) than in lots located further. This result was significant at the *p-value*=0.1 level. Among the categories of the variable pertaining to the number of adjacent unconstructed lots, only one category was significant at the 0.1 *p-value* level, and the rest were not significant. The odds of a positive lot were nearly five times higher (480%) for lots with at least a half or one adjacent unconstructed lot than for lots with no adjacent unconstructed lots (referent). This result was significant at the *p-value*=0.05 level.

All the categories of the number of canines in a lot showed significant results at least at the *p-value*=0.1 level, although the result for 4+ canines in a lot had extreme standard error. Overall a visible tendency of the odds of a lot being positive increase with the increase of the number of canines per a lot. Comparing against the referent category (one canine per lot), the odds of a lot being positive increase over three times (318%) for lots that had two canines present; and slightly more (321%) for lots that had three canines present. The lots that had four or more canines were 78 times more likely to have at least one canine positive than lots with one canine. This last result, even though it was significant at the *p-value*=0.001 level, had a very high standard error and hence its validity is limited.

5 CONCLUSIONS AND DISCUSSION

This study explored risk characteristics that may explain the prevalence of Chagas disease in canines in the Rio Grande Valley region of South Texas. The risk characteristics included those pertaining to the green open space, land development, maintenance, socio-economic status, and canines. The results of the statistical analysis are presented in Section 4, this section discusses the results in the context of previous studies and draws conclusions for policy recommendations and future research.

5.1 Interpretations of Findings – Bivariate Analysis

Among the 14 independent variables that were considered in the bivariate analysis, only four were significant at the *p=0.25* level when assessed using Fisher's exact test.

Among the Open Space variables, only Distance to Park showed a significant result (0.030 *p-value* level). The variable had two categories: Less/more than one mile distance to the nearest edge of a protected natural area, the riparian corridor of the Rio Grande and its tributaries, or a water body. One mile was selected to approximate an area of risk, because the maximum flight range of a Kissing bug is one mile (Wood &

Anderson, 1965). This results suggests that lots located closer to a natural area or a corridor may be at a higher risk of Chagas disease transmission to domestic canines.

The remaining Open Space variables: ForestALL, Uncultivated land, Pasture, and Crops did not yield statistically significant increase in the odds of the prevalence of Chagas disease, even though they provided habitat and harborage spaces for various species of wild or domesticated hosts, in addition to providing access and harborage for Kissing bugs. This result can perhaps be attributed to the overall small amounts of areas covered by these landcover types. Maximum 10% area of any sampled buffer was covered by either a forest or a pasture landcover type.

Among the Land Development variables, only Edge/ETJ was significantly associated at lower *p-value* than the cut off value (0.25) in a one-tail test. The result corroborates a prior study conducted in Mexico, where Ramirez-Sierra et al. (2010) found that peripheral location of housing within a village was a risk characteristic.

Among the Maintenance variables, only the number of unconstructed lots immediately adjacent to sampled lots was significant (*p-value*=0.039). This finding corroborates a study in central Mexico by Ramsey et al. (2005) who found that an increase in unconstructed lots increased the odds of house infestation with Kissing bugs 4.3-fold. This was the most important risk factor in their study, and was coupled with reports of roaming animals. In the context of this study, the result suggests that high rates of vacancies in peri-urban settlements may increase the risk of Chagas disease transmission to canines.

The number of canines was the most significant predictor identified in the bi-variate analysis (*p-value*=0.001). The result corroborates previous research such as that the presence of more than two dogs increases the risk of Chagas transmission (Dumonteil et al., 2013; Enger et al., 2004). This result suggest a possibility for future vector control through planning policy that could be implemented with existing pet restrictions – by limimitng the number of canines per lot.

5.2 Interpretations of Findings – Multivariate Analysis

The odds of *T. cruzi in lots* within 1-mile of a natural area (wildlife refuge, state park, riparian corridor of a river) was roughly three times higher than in lots located further. This result was significant at the *p-value*=0.089. This result is similar to the results of Ramsey et al. (2005) who found that proximity to grassland and forest were a risk factor, even though they determined the presence of grassland or forest "around the house" through surveys with householder (p.224). They also assessed the distance to river and water causeways and found no significant relationship. This result suggests that the one mile radius may be an appropriate measure for future studies of environmental risk characteristics of Chagas disease. The result also suggests that peri-urban settlements may be encroaching onto areas of wildlife habitat such as state parks, important reservoirs such as a wildlife refuge and formally unprotected natural areas such as the riparian corridor of the Rio Grande and its tributaries. The protected areas in the region are mostly located along the Rio Grande, likely forming an interconnected green refuge for many species, vectors, and pathogens.

The presence of one (or one half) unconstructed lot in the immediate adjacency increased the odds of *T. cruzi* in lots nearly twice. The result corroborates previous literature (Ramsey et al., 2005) and suggests a possibility for vector control through maintenance standards in the region. Because the high vacancy and abandonment rates are a problem across colonias (Durst & Ward, 2015), this result may hold the most relevance for vector control. Furthermore, it was noted during fieldwork that the state of overgrowth of some vacant and abandoned lots was high, in particular where overgrown lot were co-located.

The presence of more than one canine in a lot increased the odds of Chagas disease transmission with varying levels of significance (*p-value*=0.1 level for two or three canines in a lot, and p-value=0.001 for four or more canines, although with an extreme standard error). The results corroborate previous research such as the presence of more than two dogs (Dumonteil et al., 2013; Enger et al., 2004). In colonias, domestic animals such as chickens, ducks, rabbits, goats are kept as a source of subsistence or secondary income, and canines provide protection for these animals and property from predators (Erik Lee, 2013). Such living arrangements are a manifestation of disinvestment in the border rural areas, limited access to employment and municipal protections from human and animal threats. Canines in colonias, and possibly in Model Subdivisions, are kept outdoors for these reasons and as such serve as a source of bloodmeal for Kissing bugs. Some of these canines become loose and roam freely as evidence by calls from residents from unincorporated areas in Cameron county for ordinances to limit free roaming canines (The State of Texas County of Cameron, 1985). The results of the analysis offer an avenue for vector control through pet restriction policy but also offer an insight into how poverty and disinvestment can render pets a public health risk.

None of the socio-economic variables were significantly associated with the dependent variable in this study. This is surprising, however there were significant limitations to both of the variables. Both variables were allocated to individual lots from aggregated data: census block groups and county data sets. The original study by Curtis-Robles et al. (2017) found that sampling location F had significantly lower odds ratios than location A (*p-value*=0.12). The result of the current study, following the geocoding process and the sample size reduction, did not find any significant difference in odds ratio for any of the socio-economic/geographic categories. This may be explained by aggregated socio-economic data, and the sampling from an overall low-income region.

The overall model explained approximately 20% of the variance in the dependent variable and suggests an avenue for future research into environmental and land development risk characteristics.

5.3 Policy Recommendations

This section outlines several policy recommendations for implementation by the Rio Grande Valley counties.

5.3.1 Reduction and Control of Unconstructed Lots

This study corroborates previous research on the risks of unconstructed lots. This risk could be reduced through an introduction of an ordinance prescribing minimum maintenance of lots, including vacant lots, abandoned lots, and adjacent alleys. For example, the ordinance could be modeled on the public health municipal ordinance effective in the city of Brownsville in Cameron county Sec 54-62. Rubbish; weeds (Code 1971, § 16-32; Ord. No. 93-562-C, § 1, 8-31-1993) (Brownsville, 1993):

"(a) It shall be unlawful for any person who shall own any house, building, establishment, lot, yard, ground or any other place in the city or within an area immediately adjacent to the city limits and extending outside the city limits for a distance of 5,000 feet to permit or allow weeds, rubbish, brush, trash, garbage or any other matter which is unsanitary or which is unsightly or objectionable to a person of ordinary sensibilities or which is liable to produce disease to accumulate or grow thereon."

Such an ordinance could become part of the Model Subdivision rules, and be required of colonias to qualify for funding. The ordinance would seek to reduce the amount of harborage opportunities for wildlife and Kissing bugs. The larger issue of market dysfunction across colonias that results in the increase of unconstructed and abandoned lots requires the attention of policy makers, however it is beyond the scope of this thesis to suggest a policy recommendation.

5.3.2 Pet Limitation Policy

The study shows that having more than one canine per lot increases the probability of Chagas disease transmission to canines, which in turn poses a threat to human health.

Counties in the RGV should expand animal control policies in unincorporated areas that are already in place, such as the "Leash Law" in Cameron County. The ordinance (The State of Texas County of Cameron, 1985, sec. 3) prescribes the restraint of animals (dogs) by owners "to prevent such animal from running at large or attacking people or other animals or destroying the property of another person." Additionally, the ordinance requires the registration and licensing of canines that are kept in unincorporated areas (sec. 6). Both cats and dogs are prescribed to be vaccinated (sec. 7). No limit is introduced on the number of cats or canines that are permitted per one lot. Based on the results of this study, an increase in canines beyond one, increases the risk of Chagas disease, hence a recommendation to limit the number of canines per lots should be considered for unincorporated areas. Furthermore, a study to determine the effect of keeping canines in outdoor enclosure in comparison with indoor enclosures could help shape a pet control policy that would serve as a vector control mechanism and would reduce the threat of Chagas disease transmission.

An alternative approach to limiting the risks from multiple canines in lots could be based in a poverty alleviation strategy based in local investment, provision of municipal services and employment. From the literature review and in combination with the results of this study, it may be that the living arrangements of canines kept as guard dogs for other domestic animals and for security may be the risk characteristic, however future research is needed to address this assumption.

5.3.3 Protection of Wildlife and Natural Areas

At a regional scale across the counties abating the U.S. – Mexico border and the Rio Grande, an environmental assessment could be administered for the protection and preservation of natural areas, the riparian corridor of the river and its major tributaries, state parks and wildlife habitats. Following the principles of green infrastructure, patches and corridors would be connected to preserve the infrastructure function of green open spaces for both wildlife preservation, and flood mitigation. These areas could be protected by an additional buffer that would prevent the spread of vectors, such as the Kissing bug from these locations. Residential development could be restricted in this regional area for the public health benefit of local population and to preserve enough natural habitat for vectors, wildlife, and pathogens to circulate within the sylvatic cycle.

5.3.4 Limitation of Peripheral and Fragmented Development

The results of this study corroborate previous research, although not with sufficient statistical significance in the bi-variate analysis, that peripheral locations of settlements may be at higher risk of Chagas disease. If confirmed by future research studies, policies could be establish that will limit the amount of peripheries of settlements. Such policies could be based on a growth boundary limiting residential development in peri-urban areas. Alternatively, a minimum size of a subdivision, or a requirement/incentives for a contiguous development pattern of development could be established for ETJs to decrease the amount of residence exposed to sylvatic environments.

5.4 Limitations and Future Studies

This study is limited to data collected and made available through the initial study, its small size, the geographic clustering, and dependence, all of which present analytical limitations and prevent generalization from the study. These limitations were discussed in the 3.3.3 Section, previously.

Other limitations were introduced by the geocoding process and limits to the availability of secondary data, which is a common problem for studies in unincorporated areas, and small scale studies that necessitate fine grain resolution of spatial data. Measures were taken (as discussed previously in Section 3.4) to approximate the location of 13.5% of the total number of addresses. Four locations of the 104 could not be located at all. Together these make up 17.3% of the dataset, which limits its reliability. These limitations, in combination with the small clustered data set prevent a robust statistical assessment, and results of this analysis should be interpreted with caution.

5.4.1 Limited secondary data availability

This study was limited to income data aggregated into block groups. Future studies could instead utilize tax appraisal data on house values as a proxy for household socio-economic status. Studies on unincorporated regions are burdened by lack of data, such as was the case with missing tax appraisal data in this case. The incomplete county tax appraisal data sets for three of the four counties presented further limitations. Official CAD land use data were only available for Cameron county, the remaining county data sets were updated manually from aerial photos to determine the land use and vacancy rates as best as possible. This is a limitation of the study, and future studies should address this limitation.

The spatial resolution of land-cover data used in this study was 30x30 meters for each spatial cell. This resolution results in the loss of some detail as it is more appropriate for larger-scale regional analyses.

5.4.2 Fragmentation vs. green infrastructure function of landcover types

The techniques for landcover identification utilized in this study did not make a distinction about the fragmentation or clustering of land cover types, only provides information about the amount of area covered by each type. This is a limitation of the technique because it does not account for the green infrastructure function of green space features that provide patches and corridors for wildlife. For example a connected patchwork of uncultivated shrubs or forests can provide a habitat for wildlife as opposed to larger areas that are severed by cultivated or developed land, particularly infrastructure. Future studies could conduct green infrastructure analysis of patches and corridors of different types of habitat and their connections and connections with residential settlements.

REFERENCES

2013-2017 American Community Survey 5-Year Estimates. (2017a). *B19001 - Household income in the past 12 months (in 2017 inflation-adjusted dollars)*. Retrieved from: https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_17_5YR_B19001#main_content

2013-2017 American Community Survey 5-Year Estimates. (2017b). *S1903 - Median income in the past 12 months (in 2017 inflation-adjusted dollars)*. S1903. Retrieved from: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?fpt=table

Aiken, C. S. (1987). Race as a factor in municipal underbounding. *Annals of the Association of American Geographers, 77*(4), 564-579.

Anderson, M. W. (2007). Cities inside out: race, poverty, and exclusion at the urban fringe. *UCLA Law Review, 55*, 1095.

Barbu, C., Dumonteil, E., & Gourbière, S. (2010). Characterization of the dispersal of non-domiciliated Triatoma dimidiata through the selection of spatially explicit models. *PLoS Neglected Tropical Diseases, 4*(8), e777.

Barr, S. C. (2009). Canine Chagas' disease (American trypanosomiasis) in North America. *Veterinary Clinics: Small Animal Practice, 39*(6), 1055-1064.

Beard, C. B., Pye, G., Steurer, F. J., Rodriguez, R., Campman, R., Peterson, A. T., . . . Robinson, L. E. (2003). Chagas disease in a domestic transmission cycle in southern Texas, USA.

Bern, C., Kjos, S., Yabsley, M. J., & Montgomery, S. P. (2011). Trypanosoma cruzi and Chagas' Disease in the United States. *Clinical Microbiology Reviews, 24*(4), 655-681. doi:10.1128/cmr.00005-11

Bogolasky, F., & Ward, P. M. (2018). Housing, Health, and Ageing in Texas Colonias and Informal Subdivisions. *Current Urban Studies, 6*(01), 70.

Ordinances Enforcement, 1971 § Sec. 54-62. - Rubbish; weeds., 16-32 Stat. (1993 August 31, 1993).

Code of Ordinances, Appendix A. - Land Development Fees (2018 June 18, 2018).

Buehler, R. (2010). Transport Policies, Automobile Use, and Sustainable Transport: A Comparison of Germany and the United States. *Journal of Planning Education & Research, 30*(1), 76-93. doi:10.1177/0739456X10366302

Burkholder, J. E., Allison, T. C., & Kelly, V. P. (1980). Trypanosoma cruzi (Chagas)(Protozoa: Kinetoplastida) in invertebrate, reservoir, and human hosts of the lower Rio Grande valley of Texas. *The Journal of parasitology*, 305-311.

Cameron Appraisal District. (2016). *2016 - Certified Appraisal Export* [Shapefile]. 2016 - Certified Appraisal Export. Retrieved from: http://www.cameroncad.org/formsDLs_IT.html

Cantey, P. T., Stramer, S. L., Townsend, R. L., Kamel, H., Ofafa, K., Todd, C. W., . . . Dotson, E. (2012). The United States Trypanosoma cruzi Infection Study: evidence for vector‐borne transmission of the parasite that causes Chagas disease among United States blood donors. *Transfusion, 52*(9), 1922-1930.

CDC. (2013, July 19, 2013). Parasites - American Trypanosomiasis (also known as Chagas Disease) Retrieved from https://www.cdc.gov/parasites/chagas/epi.html

CDC. (Ed.) (2018) CDC A-Z Index. Center for Disease Control and Prevention.

Christensen, P., Haskins, S., Hogue, J., & Koldoff, J. (2006). Lot splitting and development regulation: the information asymmetries and free rider issues associated with Arizona's wildcat development. *Perspectives in Public Affairs, 3*, 37-59.

Cosco, N. G., Moore, R. C., & Smith, W. R. (2014). Childcare outdoor renovation as a built environment health promotion strategy: Evaluating the preventing obesity by design intervention. *American Journal of Health Promotion, 28*(SUPPL 3), S27-S32. doi:10.4278/ajhp.130430-QUAN-208

Costa, J., & Peterson, A. T. (2012). Ecological niche modeling as a tool for understanding distributions and interactions of vectors, hosts, and etiologic agents of Chagas disease. In *Recent advances on model hosts* (pp. 59-70): Springer.

Curtis-Robles, R., Zecca, I. B., Roman-Cruz, V., Carbajal, E. S., Auckland, L. D., Flores, I., . . . Hamer, S. A. (2017). Trypanosoma cruzi (Agent of Chagas Disease) in Sympatric Human and Dog Populations in "Colonias" of the Lower Rio Grande Valley of Texas. *The American journal of tropical medicine and hygiene*. doi:10.4269/ajtmh.16-0789

Davies, C. S., & Holz, R. K. (1992). Settlement evolution of 'colonias' along the US-Mexico border. *Habitat International, 16*(4), 119-142. doi:http://dx.doi.org/10.1016/0197-3975(92)90056-5

Donelson, A. J., & Esparza, A. X. (2016). *The Colonias Reader: Economy, Housing and Public Health in US-Mexico Border Colonias*: University of Arizona Press.

Dumonteil, E., Gourbière, S., Barrera-Pérez, M., Rodriguez-Félix, E., Ruiz-Piña, H., Baños-Lopez, O., . . . Rabinovich, J. E. (2002). Geographic distribution of Triatoma dimidiata and transmission dynamics of Trypanosoma cruzi in the Yucatan peninsula of Mexico. *The American journal of tropical medicine and hygiene, 67*(2), 176-183.

Dumonteil, E., Nouvellet, P., Rosecrans, K., Ramirez-Sierra, M. J., Gamboa-Leon, R., Cruz-Chan, V., . . . Gourbière, S. (2013). Eco-bio-social determinants for house infestation by non-domiciliated Triatoma dimidiata in the Yucatan Peninsula, Mexico. *PLoS Neglected Tropical Diseases, 7*(9), e2466.

Durst, N. J. (2014a). Municipal annexation and the selective underbounding of colonias in Texas' Lower Rio Grande Valley. *Environment and Planning A, 46*(7), 1699-1715.

Durst, N. J. (2014b). The rise of renters and renting in Texas colonias. *Habitat International, 43*, 72-78. doi:https://doi.org/10.1016/j.habitatint.2014.01.016

Durst, N. J. (2016). The Nature and Extent of Self-Help Housing in Texas: From Colonias to Model Subdivisions. *Journal of Planning Education and Research, 36*(2), 145-157. doi:10.1177/0739456X15612199

Durst, N. J., & Ward, P. M. (2014). Measuring self-help home improvements in Texas colonias: A ten year ‘snapshot’study. *Urban Studies, 51*(10), 2143-2159.

Durst, N. J., & Ward, P. M. (2015). Lot vacancy and property abandonment: colonias and informal subdivisions in Texas. *International Journal of Housing Policy, 15*(4), 377-399.

Durst, N. J., & Ward, P. M. (2016). Colonia Housing Conditions in Model Subdivisions: A Déjà Vu for Policy Makers. *Housing Policy Debate, 26*(2), 316-333.

Durst, N. J., & Wegmann, J. (2017). Informal Housing in the United States. *International journal of urban and regional research*.

Enger, K. S., Ordoñez, R., Wilson, M. L., & Ramsey, J. M. (2004). Evaluation of risk factors for rural infestation by Triatoma pallidipennis (Hemiptera: Triatominae), a Mexican vector of Chagas disease. *Journal of Medical Entomology, 41*(4), 760-767.

Erik Lee, C. E. W., Francisco Lara-Valencia, Carlos A. de la Parra, Rick Van Schoik, Kristofer Patron-Soberano, Eric L. Olson, Andrew Selee. (2013). The State of the Border Report. A Comprehensive Analysis of the U.S.-Mexico Border. *Border Research Partnership*.

Esteve-Gassent, M. D., de León, A. A. P., Romero-Salas, D., Feria-Arroyo, T. P., Patino, R., Castro-Arellano, I., . . . Rodriguez-Vivas, R. I. (2014). Pathogenic landscape of transboundary zoonotic diseases in the Mexico–US border along the Rio Grande. *Frontiers in public health, 2*.

Estrada-Franco, J. G., Bhatia, V., Diaz-Albiter, H., Ochoa-Garcia, L., Barbabosa, A., Vazquez-Chagoyan, J. C., . . . Garg, N. (2006). Human Trypanosoma cruzi infection and seropositivity in dogs, Mexico. *Emerging infectious diseases, 12*(4), 624.

Garcia, M. N., Aguilar, D., Gorchakov, R., Rossmann, S. N., Montgomery, S. P., Rivera, H., . . . Murray, K. O. (2015). Evidence of autochthonous Chagas disease in southeastern Texas. *The American journal of tropical medicine and hygiene, 92*(2), 325-330.

Garcia, M. N., O'Day, S., Fisher-Hoch, S., Gorchakov, R., Patino, R., Arroyo, T. P. F., . . . Jones, K. M. (2016). One Health Interactions of Chagas Disease Vectors, Canid Hosts, and Human Residents along the Texas-Mexico Border. *PLoS Negl Trop Dis, 10*(11), e0005074.

Garza, A. A. (2003). Hidalgo County. *Handbook of Texas Online.* Retrieved from http://www.tsha.utexas.edu/handbook/online/articles/view/HH/hch14.html

Garza, M., Feria Arroyo, T. P., Casillas, E. A., Sanchez-Cordero, V., Rivaldi, C.-L., & Sarkar, S. (2014). Projected Future Distributions of Vectors of Trypanosoma cruzi in North America under Climate Change Scenarios. *PLoS Neglected Tropical Diseases, 8*(5), e2818. doi:10.1371/journal.pntd.0002818

Giusti, C., & Estevez, L. (2011). Microlending for housing in the United States. A case study in colonias in Texas. *Habitat International, 35*(2), 307-315.

Gunter, S. M., Murray, K. O., Gorchakov, R., Beddard, R., Rossmann, S. N., Montgomery, S. P., . . . Widman, L. E. (2017). Likely autochthonous transmission of Trypanosoma cruzi to humans, South Central Texas, USA. *Emerging infectious diseases, 23*(3), 500.

Gürtler, R. E., & Cardinal, M. (2015). Reservoir host competence and the role of domestic and commensal hosts in the transmission of Trypanosoma cruzi. *Acta Tropica, 151*, 32-50.

Hanford, E. J., Zhan, F. B., Lu, Y., & Giordano, A. (2007). Chagas disease in Texas: recognizing the significance and implications of evidence in the literature. *Social science & medicine, 65*(1), 60-79.

Hartig, T., Mitchell, R., De Vries, S., & Frumkin, H. (2014). Nature and health. *Annual review of public health, 35*, 207-228.

Hidalgo Appraisal District. (2017). *HCAD Shapefiles* [Shapefile]. HCAD Shapefiles. Retrieved from: http://www.hidalgoad.org/maps.html

Hodo, C. L., & Hamer, S. A. (2017). Toward an Ecological Framework for Assessing Reservoirs of Vector-Borne Pathogens: Wildlife Reservoirs of Trypanosoma cruzi across the Southern United States. *ILAR Journal, 58*(3), 379-392. doi:10.1093/ilar/ilx020

Hotez, P. J., Bottazzi, M. E., Dumonteil, E., Valenzuela, J. G., Kamhawi, S., Ortega, J., . . . Tapia-Conyer, R. (2012). Texas and Mexico: Sharing a Legacy of Poverty and Neglected Tropical Diseases. *PLoS Neglected Tropical Diseases, 6*(3), e1497. doi:10.1371/journal.pntd.0001497

Kassambara, A. (2018). Logistic Regression Assumptions and Diagnostics in R. Retrieved from http://www.sthda.com/english/articles/36-classification-methods-essentials/148-logistic-regression-assumptions-and-diagnostics-in-r/

Kjos, S. A., Marcet, P. L., Yabsley, M. J., Kitron, U., Snowden, K. F., Logan, K. S., . . . Dotson, E. M. (2013). Identification of Bloodmeal Sources andTrypanosoma cruziInfection in Triatomine Bugs (Hemiptera: Reduviidae) From Residential Settings in Texas, the United States. *Journal of Medical Entomology, 50*(5), 1126-1139. doi:10.1603/me12242

Kjos, S. A., Snowden, K. F., & Olson, J. K. (2009). Biogeography and Trypanosoma cruzi infection prevalence of Chagas disease vectors in Texas, USA. *Vector Borne Zoonotic Dis, 9*(1), 41-50. doi:10.1089/vbz.2008.0026

Klotz, S. A., Shirazi, F. M., Boesen, K., Beatty, N. L., Dorn, P. L., Smith, S., & Schmidt, J. O. (2016). Kissing bug (Triatoma spp.) intrusion into homes: Troublesome bites and domiciliation. *Environmental health insights, 10*, EHI. S32834.

Krellenberg, K., Welz, J., & Reyes-Päcke, S. (2014). Urban green areas and their potential for social interaction – A case study of a socio-economically mixed neighbourhood in Santiago de Chile. *Habitat International, 44*, 11-21. doi:https://doi.org/10.1016/j.habitatint.2014.04.004

Lichter, D. T., Parisi, D., Grice, S. M., & Taquino, M. (2007). Municipal underbounding: Annexation and racial exclusion in small southern towns. *Rural Sociology, 72*(1), 47-68.

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., . . . Nieuwenhuijsen, M. J. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research, 158*, 301-317.

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. Thousand Oaks, CA: SAGE Publications, Inc.

National Land Cover Database 2011 (2011). Retrieved from: https://www.mrlc.gov/data

NIEHS. (2017, July 20, 2017). Vectorborne and Zoonotic Diseases. Retrieved from https://www.niehs.nih.gov/research/programs/geh/climatechange/health_impacts/vectorborne/index.cfm

Nouvellet, P., Cucunubá, Z. M., & Gourbière, S. (2015). Ecology, evolution and control of Chagas disease: a century of neglected modelling and a promising future. In *Advances in parasitology* (Vol. 87, pp. 135-191): Elsevier.

Pacheco-Tucuch, F. S., Ramirez-Sierra, M. J., Gourbière, S., & Dumonteil, E. (2012). Public street lights increase house infestation by the Chagas disease vector Triatoma dimidiata. *PLOS ONE, 7*(4), e36207.

Packchanian, A. (1939). Natural infection of Triatoma gerstakeri with Trypanosoma cruzi in Texas. *Public Health Reports (1896-1970)*, 1547-1554.

Packchanian, A. (1942). Reservoir hosts of Chagas' disease in the state of Texas. *The American journal of tropical medicine and hygiene, 1*(6), 623-631.

Parcher, J. W., & Humberson, D. (2007). *Colonia boundaries in Texas* [Shapefile]. Retrieved from: https://pubs.usgs.gov/of/2007/1230/

Peterson Townsend, A., Soberon, J., Pearson, R. G., Anderson, R. P., Mertinez-Meyer, E., Nakamura, M., & Bastos Araujo, M. (2011). *Ecological niches and geographic distributions* Princeton, NJ: Princeton University Press.

Pew Research Center. (2016). Hispanic Population Growth and Dispersion Across U.S. Counties, 1980-2014. Retrieved from http://www.pewhispanic.org/interactives/hispanic-population-by-county/

Raghavan, R. K., Saunders, A. B., Goodin, D. G., Anderson, G. A., & Harkin, K. R. (2015). Geospatial Risk Factors of Canine American Trypanosomiasis (Chagas Disease)(42 Cases: 2000–2012). *Vector-Borne and Zoonotic Diseases, 15*(10), 602-610.

Ramírez, J. D., Turriago, B., Tapia-Calle, G., & Guhl, F. (2013a). Understanding the role of dogs (Canis lupus familiaris) in the transmission dynamics of Trypanosoma cruzi genotypes in Colombia. *Veterinary parasitology, 196*(1-2), 216-219.

Ramírez, J. D., Turriago, B., Tapia-Calle, G., & Guhl, F. (2013b). Understanding the role of dogs (Canis lupus familiaris) in the transmission dynamics of Trypanosoma cruzi genotypes in Colombia. *Veterinary parasitology, 196*(1), 216-219.

Ramirez ‐ Sierra, M. J., Herrera ‐ Aguilar, M., Gourbière, S., & Dumonteil, E. (2010). Patterns of house infestation dynamics by non ‐ domiciliated Triatoma dimidiata reveal a spatial gradient of infestation in rural villages and potential insect manipulation by Trypanosoma cruzi. *Tropical Medicine & International Health, 15*(1), 77-86.

Ramsey, J., Alvear, A., Ordonez, R., Munoz, G., Garcia, A., Lopez, R., & Leyva, R. (2005). Risk factors associated with house infestation by the Chagas disease vector Triatoma pallidipennis in Cuernavaca metropolitan area, Mexico. *Medical and veterinary entomology, 19*(2), 219-228.

Ramsey, J., Ordoñez, R., Cruz‐Celis, A., Alvear, A., Chavez, V., Lopez, R., . . . Carrillo, S. (2000). Distribution of domestic Triatominae and stratification of Chagas disease transmission in Oaxaca, Mexico. *Medical and veterinary entomology, 14*(1), 19-30.

Ramsey, J. M., Peterson, A. T., Carmona-Castro, O., Moo-Llanes, D. A., Nakazawa, Y., Butrick, M., . . . Ibarra-Cerdeña, C. N. (2015). Atlas of Mexican Triatominae (Reduviidae: Hemiptera) and vector transmission of Chagas disease. *Memórias do Instituto Oswaldo Cruz, 110*(3), 339-352.

Ryabov, I., & Merino, S. (2017). Recent Demographic Change in the Rio Grande Valley of Texas: The Importance of Domestic Migration. *Journal of Borderlands Studies, 32*(2), 211-231.

Ryckman, R. E. (1981). The Kissing Bug Problem In Western North America. *Bulletin of the Society of Vector Ecologists, 6*, 167–169.

Sarkar, S., Strutz, S. E., Frank, D. M., Rivaldi, C. L., Sissel, B., & Sánchez–Cordero, V. (2010). Chagas disease risk in Texas. *PLoS Negl Trop Dis, 4*(10), e836.

Shanahan, D. F., Lin, B. B., Bush, R., Gaston, K. J., Dean, J. H., Barber, E., & Fuller, R. A. (2015). Toward improved public health outcomes from urban nature. *American Journal of Public Health (ajph)*.

Shih, W. (2017). Greenspace patterns and the mitigation of land surface temperature in Taipei metropolis. *Habitat International, 60*, 69-80. doi:https://doi.org/10.1016/j.habitatint.2016.12.006

Shultz, M. M., & Groy, J. B. (1988). The failure of subdivision control in the Western United States: A blueprint for local government action. *Utah L. Rev.*, 569.

Starr Appraisal District. (2017). *SCAD Parcels* [Shapefile]. SCAD_PARCELS_8142017.

State Park Boundaries. (n.d.). Retrieved from: https://tpwd.texas.gov/gis/#gis

Tenney, T. D., Curtis-Robles, R., Snowden, K. F., & Hamer, S. A. (2014). Shelter dogs as sentinels for Trypanosoma cruzi transmission across Texas. *Emerging infectious diseases, 20*(8), 1323.

Texas Comptroller of Public Accounts. (2014). *Texas Property Tax Assistance Property Classification Guide*. Retrieved from

Texas Orthoimagery. (2016). Aerial map. Retrieved June 7, 2018, from Texas Orthoimagery Program http://www.arcgis.com/home/webmap/viewer.html?webmap=c1c2090ed8594e0193194b750d0d5f83

The Colonia Initiatives Program. (2006). *Tracking the Progress of State-funded Projects that Benefit Colonias*. Retrieved from

Ordinance #1985-1, (1985).

U.S. Census Bureau. (2016). *2017 Tiger/Line Shapefiles* [Shapefiles]. cb_2016_48_place_500k, tl_2016_48_tract, tl_2016_us_county, tl_2016_us_state. Retrieved from: https://www.census.gov/cgi-bin/geo/shapefiles/index.php

U.S. Census Bureau. (2017). *Poverty thresholds by Size of Family and Number of Children* [Excel]. thresh17.xls. Retrieved from: https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html

U.S. Geological Survey. (2014). *NLCD 2011 Land Cover* [remote-sensing image]. National Geospatial Data Asset (NGDA) Land Use Land Cover. Retrieved from: http://www.mrlc.gov

UCLA Institute for Digital Research and Education. (n.d.). Lesson 3 Logistic Regression Diagnostics. In Logistic Regression with Stata. Web.: UCLA Institute for Digital Research and Education. Retrieved from https://stats.idre.ucla.edu/stata/webbooks/logistic/chapter3/lesson-3-logistic-regression-diagnostics-2/.

Ward, P. M. (1999). *Colonias and public policy in Texas and Mexico: Urbanization by stealth*. Austin: University of Texas Press.

Ward, P. M., & Carew, J. (2000). Absentee lot owners in Texas colonias: who are they, and what do they want? *Habitat International, 24*(3), 327-345. doi:https://doi.org/10.1016/S0197-3975(99)00047-8

Ward, P. M., de Souza, F., & Giusti, C. (2004). 'Colonia' Land and Housing Market Performance and the Impact of Lot Title Regularisation in Texas. *Urban Studies, 41*(13), 2621-2646.

Ward, P. M., & Peters, P. A. (2007). Self-help housing and informal homesteading in peri-urban America: Settlement identification using digital imagery and GIS. *Habitat International, 31*(2), 205-218.

WHO Expert Committee. (2002). Control of Chagas disease. *World Health Organization technical report series, 905*, i.

Wildlife Management areas (n.d.). [GIS]. Retrieved from: https://tpwd.texas.gov/gis/#gis

Willacy Appraisal District. (2017). *Willacy CAD Shapefiles* [Shapefile]. Willacy CAD Shapefiles.

Williams, G. D., Adams, L. G., Yaeger, R. G., McGrath, R. K., Read, W. K., & Bilderback, W. R. (1977). Naturally occurring trypanosomiasis (Chagas' disease) in dogs. *Journal of the American Veterinary Medical Association, 171*(2), 171-177.

Williams, R. (2018). *Scalar Measures of Fit: Pseudo R2 and Information Measures (AIC & BIC)*. Retrieved from Web: https://www3.nd.edu/~rwilliam/stats3/L05.pdf

Wood, S. F., & Anderson, R. C. (1965). Conenose Bugs (Triatoma) Visit Unoccupied Boy's Camp in Los Angeles. *Journal of Medical Entomology, 1*(4), 347-348. doi:10.1093/jmedent/1.4.347

World Health Organization. (2016). *Urban green spaces and health - a review of evidence*. Retrieved from http://www.euro.who.int/en/health-topics/environment-and-health/urban-health/publications/2016/urban-green-spaces-and-health-a-review-of-evidence-2016

Wozniak, E. J., Lawrence, G., Gorchakov, R., Alamgir, H., Dotson, E., Sissel, B., . . . Murray, K. O. (2015). The biology of the triatomine bugs native to south central Texas and assessment of the risk they pose for autochthonous Chagas disease exposure. *Journal of Parasitology, 101*(5), 520-528.

Xu, L., You, H., Li, D., & Yu, K. (2016). Urban green spaces, their spatial pattern, and ecosystem service value: The case of Beijing. *Habitat International, 56*, 84-95. doi:https://doi.org/10.1016/j.habitatint.2016.04.005

Zeledon, R., Beard, C. B., & Dias, J. P. (2012). *An appraisal of the status of Chagas disease in the United States*: Elsevier.